



273	Db	GGCGGACCTCTGCGCTTTTCGCTTACGACATCTTTGACAAACAGCATTTTCAGTACGGCAACAGG	332
300	Qy	GTATTTGCGCAATACCCAGAGACATAGCAGATTTATTTCAAGCAGACGTTTCTCTGAGGGG	359
333	Db	GTATTCACCAATACCCAAAGACATACGAGCTATTTCAAGCAGTCGTTTCTGAGGGG	392
360	Qy	TACTTCTGGGAACGAAGCATGACATACGAAGACACGAGGCATTTTGGATCGCCACAAACGCAC	419
393	Db	TATTCTCTGGGAAGAAGCATGACTTTTCGAAGACACGAGGCGTTTGCACCGTCAAGCGAC	452
420	Qy	ATAACATGATGGGAAGGCGTCGACGCACTGTTTTCCTATATAAATTCGATTTTCATGGTGTG	479
453	Db	ATAAAGTTG-----GAAGGCGCACTGTTTTTTTCTACGAAATTCGATTTTATGGTGTG	503
480	Qy	AACTTTCTCGCAATGGTCCAGTTATGCAGAGGAAGACGCTGAAATGGGAGGCATCCACT	539
504	Db	AACTTTCCCTCAAGTGGTCCAGTTATGCAGAAGAAGACGCTGAAATGGGAGGCATCCACT	563
540	Qy	GAGATAATGTATGCGCGTGATGGAGTGTCTGAAGGGTGATTTAAACATGCGCTCTGTTCGTT	599
564	Db	GAGAAATGTACGTGCGTGATGGAGTGTCTACTGGGGGATGTTAGCAGGACGCTGTGGCTT	623
600	Qy	GAAGGAGTGGCCATTACCGATGTGACTTCAAAACTATCTTACAAAGCTAAGAAGGTTGTGC	659
624	Db	GAAGGGATAAACATCACCGATGTAACTTCAGAACTTACCGGGCAAGAAGGGTGTGC	683
660	Qy	CGGTTGCCAGACTATCATTTTGTGGACCATCGCATTTGAGATGTGCGGCCACACCAAGAAT	719
684	Db	GTGTTGCCAGAAATATCATTTTGTGTGGACCAACCCGAATTGAAATTTCTGAGGCCATGACAAAGAT	743
720	Qy	TACAAACAGGTTTAAGCTGCACGAGCATGCCGAAGCTCG---TCATGGACTTGTCAAGGAAG	776
744	Db	TACAAACCCGTTGAGGTGTATGAGAAATGCCGTTGCTCGCCCTTCTATGCTGCGGGTTAAG	803
777	Qy	GCCAAGTAAAGCCTTAATGAAAAAGTCAAGACGAACACGAGGAGAAACAAA	826
804	Db	GCCAAGTAAAGGCTTAACTTAAAGGCCCAACGAACACGAGGAGGAAAAAA	853

RESULT 2  
US-09-459-956-7

```

; Patent NO. 0342379
;
; GENERAL INFORMATION:
;
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, Iii, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7

```

Query Match 39.4%; Score 419.6; DB 3; Length 801;  
Best Local Similarity 75.6%; Ref. No. 1.3e-127;  
Matches 537; Conservative 0; Mismatches 164; Indels 9; Gaps 1;  
QY 52 AGGTTGATATCTTACTTACGTTCTACCATCATGACAAAGTTTGCACAGGAAAAGGCTGTGA 111  
Db 77 AGGCTGACTTAGAAGAGAAGACATTGAGATATACAAAAGCTCTAACACCACTGGGTGTGA 136

Qy	112	TTAAAC	CAGACATGAAGATGAAGCTGCGTATGGAAGGTGCTGTATAAACCGGCAACAAGTTCCG	171
Db	137	TTAAAC	CAGACATGAAGATTAAGCTGAAGATGGAAGGAAATGATAAACCGGCGATGCTTTTG	196
Qy	172	TGTTGA	AGGAGATGGAATAAGGGGAAGCCTTTTCGACGGAAACACAGACTATGGACCTTACAG	231
Db	197	TGATCG	AAGGAGGAGGAGGAAGGAAAGCCTTACGATGGGACACACACTTTAAACCTTGGAG	256
Qy	232	TCATAGA	AGGCGCACCATTCGCTTTTCGCTTACGATATCTTGACCAACAGTATTCGATTAAG	291
Db	257	TGAAGGA	AGGTGCGCCTCTGCTTTTCTTACGATATCTTGTCAAAACGCGTTCACGATACG	316
Qy	292	GCAACAG	GGGTATTCGCCAAATACCCAGAGACATAGCAGATTATTTCAAGCAGACGTTTC	351
Db	317	GAACAC	AGACATGACAAAATACCCAGACGATATAGCAGACTATTTCAAGCAGTCTGTTTC	376
Qy	352	CTGAGGG	GTACTCTTCGGGAACGAAGCATGACATACGAAGACCAAGGCGCATTTGCATCGCCA	411
Db	377	CCGAGG	GATATTCCTGGGAAGAACCATGACTTTTGAAGACAAAGGCATTTGTCAAGTGA	436
Qy	412	CAAAAG	CATAAACAATGATGGAAGGCTGACGACTGTTTTCCTATATAAAATTCGATTTG	471
Db	437	AAAGTGA	TCATAAGCATGAGGAA-----GACTCCTTTATCTATGAAATTCGTTTTC	487
Qy	472	ATGCTGT	GAACCTTTTCCTGCCAATCGTCCAGTTATGCGAGAGGAAGCGCTGAAATGGGAGC	531
Db	488	ATGGGAT	GAACCTTTTCCTCCCAATGGTTCGGTTATGCGAAAAAAACCTTTGAAGTGGGAAC	547
Qy	532	CATCCAC	TGAGATAATATATATGCGCGTGATGGAGTGTCTGAAGGGTGATGTTAACTGGCTC	591
Db	548	CATCCA	CTGAGATTATGTAATGCGGTGATGGAGTGTCTGGTCGAGATATTAGGCCATTCTC	607
Qy	592	TGTTGCT	TGAGAGGTGGCCATTACCGATGTGACTTCAAAACCTATTACAAGCTAAGA	651
Db	608	TGTTGCT	TGGAGGGAGGTGGCCATTACCGATGTGACTTCAAAAGTATTTACAAGGCAGAAA	667
Qy	652	AGGTTGT	CCGCTGCCAGACTATCACTTTTGTGGACCATCGCATTTGAGATTGTGAGCCACG	711
Db	668	AAGTTGT	CAAAATGCCAGACTATCACTTTTGTGGACCATCGCATTTGAGACTCTTGAACCATG	727
Qy	712	ACAAAGA	ATTACAACAGGTTAAGCTGCACGAGCATGCGGAAGCTCGTCAT	761
Db	728	ACRAGGA	TTACAACAAAGTAACGCTGTATGAGAAATGCAAGTTCGCTGCTAT	777

```

RESULT 3
US-09-459-956-6
; Sequence 6, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp
US-09-459-956-6

```

Query Match	24.1%;	Score 256.6;	DB 3;	Length 678;
Best Local Similarity	62.8%;	Pred. No. 7.4e-74;		







QY	702	GTGAGCCACGCAAGAATTACAAAGGTTAAGCTGTCACGAGCATGCCGAAGCTCGTCAT	761
Db	3577	ACCTCCCAACACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGCCAC	3636
QY	762	GGACTGT	768
Db	3637	CACCTGT	3643

RESULT 9  
 US-09-459-956-2  
 ; Sequence 2, Application US/09459956  
 ; Patent No. 6342379  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tsien, Roger Y.  
 ; APPLICANT: Gonzalez, III, Jesus E.  
 ; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
 ; TITLE OF INVENTION: OPTICAL METHODS  
 ; FILE REFERENCE: REGEN1290-4  
 ; CURRENT APPLICATION NUMBER: US/09/459,956  
 ; PRIOR FILING DATE: 1999-12-13  
 ; PRIOR APPLICATION NUMBER: 08/765,860  
 ; PRIOR FILING DATE: 1999-05-08  
 ; PRIOR APPLICATION NUMBER: 08/481,977  
 ; PRIOR FILING DATE: 1995-06-07  
 ; PRIOR APPLICATION NUMBER: PCT/US96/09652  
 ; PRIOR FILING DATE: 1996-06-06  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 690  
 ; TYPE: DNA  
 ; ORGANISM: Anemonia majano  
 US-09-459-956-2

QY 511 GGAAGACGCTGAATGGAGCCATCCACTGAGATAATGTATGCGCGTAT-----GGAG 564  
Db 419 AGATGACAACTAATCTGGAGCATCTCGGAGAGATCATGCGAGTACCTTAACGAGGGA 478  
QY 565 TGCTGAAGGCTGATGTTAAATCATGGCTCTGTTGCTTGAAGAGAGTGCCCATTTACCGATGTG 624  
Db 479 TACTGAAGGGGATGTCTCCATGATCTCTCTCTGAAGGATGCTGGCGGTTACCGGTGCC 538  
QY 625 ACTTCAAACTACTTACAAAGCTAAGAGGT 655  
Db 539 AGTTCGACACAGTTTACAAAGCAAAAGTCTGT 569

## RESULT 11

US-09-459-956-3  
; Sequence 3, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, Iii, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; FILE REFERENCE: OPTICAL METHODS  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 696  
; TYPE: DNA  
; ORGANISM: Zoanthus sp  
US-09-459-956-3

Query Match 16.5%; Score 176.2; DB 3; Length 696;  
Best Local Similarity 59.4%; Pred. No. 2.6e-47;  
Matches 339; Conservative 0; Mismatches 223; Indels 9; Gaps 2;

QY 91 TTGCACAGAAAGGCTGATTAAACACAGACATGAAGATGAAGCTGCGTATGGAGGTG 150  
Db 2 TGCGCTAGTCAAGCAGCGTCTAAACAAAGAAATGACATGAATACCGTATGGAGGTG 61  
QY 151 CTGTAACCGGGCACAGTTCTGCTGTTGAAGGAGATGGAAGGAGCGCTTTTCGACGGAA 210  
Db 62 GCGTCGATGACATAAATTTGTATCACCGGAGAGGCGATTGGATATCGTTCAAGGGA 121  
QY 211 CACAGACTATGACCTTACAGTCATAGAGGCGACCATTTGCTTTCGCTTACGATATCT 270  
Db 122 AACAGCTATTAATCTGTGTGTGTCGAAGGTGGACCATTTGCGGAAGACATAT 181  
QY 271 TGACAACTATTCGATTAAGGCAACAGGCTATTCGCAAAATACCCAGAGACATAGCAG 330  
Db 182 TGTACGCTCCCTTTTACCTACGGAACAGGGTTTTCACCTGAATATCTCTAAGACATAGTTG 241  
QY 331 ATTATTTCAAGCAGAGCTTTCTCTGAGGGTACTTCTGGGAACGAAGCATGACATACGAAG 390  
Db 242 ACTATTTCAAGAACTGCTGCTCTGCTGGATATACATGGAGAGGCTTTTCTCTTTGAGG 301  
QY 391 ACCAGGGCATTTGCATCGCCACAAACGACATAAACATATGGAAGCGGTCGACGACTGTT 450  
Db 302 ATGGAGCAGTTTGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 358  
QY 451 TTGCCTATAAAATTCGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 510  
Db 359 TGTATCATGAGTCCAAATTTTATGAGTGAATTTTCTCTGCTGATGACCTGTGATGAAAA 418  
QY 511 GGAAGACGCTGAATGGAGCCATCCACTGAGATAATGTATGCGCGTAT-----GGAG 564

Db 419 AGATGACAGATAAATCTGGAGCCATCTCGAGAGATCATACGAGTACCTTAACGAGGGA 478  
QY 565 TGCTGAAGGCTGATGTTAAATCATGGCTCTGTTGCTTGAAGAGAGTGCCCATTTACCGATGTG 624  
Db 479 TATTGAAGGGGATGTCTCCATGATCTCTCTCTGAAGGATGCTGGCGGTTACCGGTGCC 538  
QY 625 ACTTCAAACTACTTACAAAGCTAAGAGGT 655  
Db 539 AATTGACACAGTTTACAAAGCAAAAGTCTGT 569

## RESULT 12

US-09-977-897-1  
; Sequence 1, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresc  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems a  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1482  
; TYPE: DNA  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-1

Query Match 13.1%; Score 139.2; DB 3; Length 1482;  
Best Local Similarity 51.5%; Pred. No. 6.8e-35;  
Matches 352; Conservative 0; Mismatches 323; Indels 9; Gaps 1;

QY 73 CTACCATCATGACAAAGTGTTCACAGAGAAAGGGTGTGATTAATTAACACAGACATGAAGATGA 132  
Db 8 CCACCATGTTGACCGGAAACGCTGTGAAGAACACCGGCTGAGAGAGATCATGAGGCCA 67  
QY 133 AGCTGCTGATGAAGGCTGCTGTAAACGGGCAAGAGTTCGTGTTGAAAGAGATGAAAG 192  
Db 68 AGGCCAGCGTGGAGGGCATCTGTAACCAACACCGTGTTCAGCATGAGGGCTTCGGCAAG 127  
QY 193 GGAAGCCTTTTCGAGGACACAGACTATGGACCTTACAGTATAGAAAGGCGCACCATTCG 252  
Db 128 GCAACGCTGCTGTTGCGCAACCGAGCTGATGAGATCCGGGTGACCAAGGCGGCCCTCTGC 187  
QY 253 CTTTCGCTTACGATATCTTGACAAAGTATTCGATTCAGGCAACAGGGTATTCGCCAAAT 312  
Db 198 CTTTCGCTTTCGACATCTGTGAGCATCGCTTCCAGTACGGCNAACGGACCTTCACCAAGT 247  
QY 313 ACCAGAAAGACATAGCAGATTTTCAAGCAGAGCTTTCTGAGGGGTACTTCTGGGAAC 372  
Db 248 ATCCGACGACATCGCCGACTTCTGTCGAGAGCTTCCCTGCGGCTTCTTCTACGAGC 307  
QY 373 GAAGCATGACATAGAGACACAGGCGCATTTGCAATGCGCACAAACGACATACAAATGATGG 432  
Db 308 GGAACCTGCGGTTTCGAGGACCGGCGCCATCTGTSBACATCCGGAGCGACATCAGCCTG 363  
QY 433 AAGGCTGCGAGCATGTTTTCCTTATATAATTTTCGATTTGATGTTGTTGAACTTTCCTGCCA 492  
Db 364 -----GAGGACGACAGTTCCACTACAGGTGGAGTACCGGCGGCAACGGCTTCCCTAGCA 418  
QY 493 ATGTTCCAGTTTATGAGAGGAGACGCTGAAATGGAGGCCATCCACTGAGATAATATGATG 552  
Db 419 ACGGCTTGTGATGACAGAGGCCATCTCTGGCATGGAGCCCGAGCTTCGAGGTGGTGTACA 478  
QY 553 CGGCTGATGAGTCTGAGAGGCTGATGTTAAATGATGGCTCTGTTGCTTGAAGAGGTGGCC 612  
Db 479 TGAACAGCGGCGTCTGTTGGGCGAGGTGGACCTTGGTGTACAAGCTGAGAGCGGCAACT 538  
QY 613 ATTACCGATGTGACTTCAAACTACTTACAAAGCTTACAAAGCTAGAAAGGTTGTCGGTTGCCAGCT 672

Db 539 ACTACAGTCGCCACATGAAGACCTTCTACCGAGCAAGGGCGCGTGAAGGAGTTCCCTG 598  
Qy 673 ATCACTTTGTGACCATCGCATTCAGATTGTGAGCCAGCAAAAGATTACAAAGGTTA 732  
Db 599 AGTACCTATTTCATCCACCACCGGTGGAGAAAGAACTACGTGGAGGAGGGCAGCTTCGTG 658  
Qy 733 AGCTGCACGACATGCGCGAAGCTC 756  
Db 659 AGCAGCAGCAGACCGCCATCGCCC 682

## RESULT 13

US-10-021-818A-1  
; Sequence 1, Application US/10021818A  
; Patent No. 6936428  
; GENERAL INFORMATION:  
; APPLICANT: Davis, Ronald W.  
; APPLICANT: Vaillancourt, Peter  
; TITLE OF INVENTION: Dimeric Fluorescent Polypeptides  
; FILE REFERENCE: 25436/1652  
; CURRENT APPLICATION NUMBER: US/10/021,818A  
; CURRENT FILING DATE: 2001-12-13  
; PRIOR APPLICATION NUMBER: US 60/256,121  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Renilla reniformis  
US-10-021-818A-1

Query Match 12.4%; Score 132.2; DB 3; Length 720;  
Best Local Similarity 54.4%; Pred. No. 8.7e-33;  
Matches 316; Conservative 0; Mismatches 253; Indels 12; Gaps 2;  
Qy 122 CATGAAGATGAAGCTGCGTATGGAGGTGCTGTAACGGGCAACAGTTTCGTGGTTGAAGG 181  
Db 45 CATGTCGTTTAAGTGATCTTGGAGGTGTAGTAACAAATCATGTGTTCACAAATGGAAGG 104  
Qy 182 AGATGGAAAGGAGGCTTTTCGACGGAAACACAGACTATGACCTTACAGTCATAGAAGG 241  
Db 105 TTGTGGAAAGGAAATATTTATTTCGGAACCAACTGGTTTCAGATTTCGTGTACAAAAGG 164  
Qy 242 CGCACCATTCGCTTCGTATGAGTATCTTCACACAGATTCATTCAGGCACAGGGT 301  
Db 165 GGTCCCGCTTCCATTTGATTTGATATTTCTCTCACCGACTTTCCAATACGGCAACCGTAC 224  
Qy 302 ATTCGCCAAATACCCAGAGACATAGCAGATTATTTCAAGCAGACGTTTCTCGAGGGGTA 361  
Db 225 ATTCACGAAATACCCGAGGATATATCAGACTTTTATATACATCAATTTCCAGCGGAT 284  
Qy 362 CTTCTGGGAACCAAGCATGACATACGAAGACACGAGGCAATTTGCATCCGCCAACAACGACAT 421  
Db 285 TGTATACGAAGAAGCTTGGCTTACGAAGATGGTGAGCTGGTTGAAATCCCGTTCAGATAT 344  
Qy 422 AACATGATGAAGCGTCGAGGACTGTTTTCCTATTAATTCGATTTGATGCTGTGA 481  
Db 345 AAATTTAA-----TCGAGGAGATGTTTGTCTACAGAGTGGAATATAAAGGTAGTAA 395  
Qy 482 CTTTCTGCCAATGGTCAGATTATGACAGGAAGACGCTGAAATGGGAGCCATCCACTGA 541  
Db 396 CTTCCCGAATGATGTCAGTGATGAAGAAGACATACAGGATTCACACCTTCGTTTGA 455  
Qy 542 GATTAATGATCGCGTGTGAGGTGCTGAAGGGTGAATTTAAACATGGCTCTGTTGCTTGA 601  
Db 456 AGTTGTGATATGAACGATGCGCTCTTGTGGCCAAAGTCAATCTTGTATTATAGATTA 515  
Qy 602 AGGAGGTGGCCATTACCGATGTGACTTCAAACTACTTCAAAAGCTAAGAAGGTGTCGG 661  
Db 516 CTCTGGCAAAATTTTATTCGTGTGTCATGAGAACAACCTGATGAATCAAAAGGGTGTAGTAA 575  
Qy 662 G---TTGCCGACATATCACTTTGTGGACCACCGCTTGGAAGAACCTACGTGGAGACGGC 699

Db 576 GGATTTTCCGATACCAATTTTCATTCAACATCGTTTAGAGA 616

## RESULT 14

US-09-839-650-1  
; Sequence 1, Application US/09839650  
; Patent No. 6645761  
; GENERAL INFORMATION:  
; APPLICANT: Stratgene  
; TITLE OF INVENTION: Humanized Polynucleotide Sequence Encoding Renilla Mulleri Green  
; TITLE OF INVENTION: Fluorescent Protein  
; FILE REFERENCE: 25436/1755  
; CURRENT APPLICATION NUMBER: US/09/839,650  
; CURRENT FILING DATE: 2001-04-19  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanized R. mulleri polynucleotide  
; NAME/KEY: misc feature  
; LOCATION: (1)..(720)  
; OTHER INFORMATION: Humanized DNA sequence  
US-09-839-650-1

Query Match 11.8%; Score 126; DB 3; Length 720;  
Best Local Similarity 51.3%; Pred. No. 9.7e-31;  
Matches 325; Conservative 0; Mismatches 300; Indels 9; Gaps 1;  
Qy 123 ATGAAGATGAAGCTGCGTATGGAGGTGCTGTAACGGGCAACAGTTTCGTGGTTGAAGG 182  
Db 46 ATGAGCTTACAAGGTGAACCTTGGAGGGCATCGTGAACCAACCACTGTTTCACCATGAGGGC 105  
Qy 183 GATGGAAGGAGGAGCCCTTTCGACGGAAACACAGACTATGACCTTACAGTCATAGAAGGC 242  
Db 106 TCGGGCAAGGCAACATCCTGTTTCGGCAACCACTGCTGTGCAGATCCGGTGAACCAAGGC 165  
Qy 243 GCACCATTCGCTTTCGCTTACGATATCTTGACAAACAGTATTCGATTAACGGCAACAGGTA 302  
Db 166 GCCCCCTCGCTTCGCTTCGACATCGTGAGCCCGCTTCCAGTACGGCAACCGCAC 225  
Qy 303 TTGCCAAATACCCAGNAGACATAGCAGATTATTTCAAGCAGACGTTTCTCGAGGGGTAC 362  
Db 226 TTCACCAAGTACCCCAACGACATCAGCGACTACTTTCATCCAGAGCTTCCCGCGCGCTTC 285  
Qy 363 TTCTGGGAACGAAGCATGACATACGAAGACCAAGGCAATTTGCATCGCCCAACAACGACATA 422  
Db 286 ATGTACGAGCGCACCTCGCTACGAGGACGGCGCTCGTGGAGATCCCGCAGGACATC 345  
Qy 423 ACAATGATGAAGGCGTCGACGACTGTTTTCCTATAAATTTGATTTGATGTTGTAAC 482  
Db 346 AACCTGA-----TCGAGGACAAAGTTCTGTGTACCGCTGGAGTACAAGGGCAGCAAC 396  
Qy 483 TTTCCTGCCAATCGTCCAGTTATGACAGGAAGACGCTGAAATGGGAGCCATCCACTGAG 542  
Db 397 TTCCCGACACCGGCCCGGCTGATGACAGAACCATCTCTGGGCATCGAGCCAGCTTCGAG 456  
Qy 543 ATAATGATCGCGTGTATGAGTGTGTAAGGGTGAATTTAAACATGGCTCTGTTCTTGA 602  
Db 457 GCATGTGATGAACAAACGCGTGTGTTGGCGAGGTGATCTCTGGTGTACAAGCTGAAC 516  
Qy 603 GGAGGTGGCCATTACCGATGTGACTTCAAACTACTTCAAAAGCTAAGAAGGTGTCGG 662  
Db 517 AGCGGCAAGTACTACAGCTGCCACATGAACACCTGATGAAGAGCAAGGGCGTGTGAAG 576  
Qy 663 TTGCCGACATCACTTTTGTGGACCATCGCATTCGATTCGAGATTGTGAGCCACGACAAAGATTAC 722  
Db 577 GAGTTCCCTCTTACCACTTTCATCCAGCACCGCTTGGAAGAACCTACGTGGAGACGGC 636



QY 723 AACAGGTTAAGCTGACGAGCATGCCGAAGCTC 756  
| | | | | | | | | | | | | | | | | | | | | |  
Db 637 GGCTTCGTGGAGCAGCAGCAGCCGCGCATCGCCC 670

## RESULT 15

US-10-652-703A-1  
; Sequence 1, Application US/10652703A  
; Patent No. 6884620  
; GENERAL INFORMATION:  
; APPLICANT: Stratagene  
; TITLE OF INVENTION: Humanized Polynucleotide Sequence Encoding Renilla Mulleri Green  
; Patent No. 6884620  
; TITLE OF INVENTION: Fluorescent Protein  
; FILE REFERENCE: 25436/1754  
; CURRENT APPLICATION NUMBER: US/10/652,703A  
; CURRENT FILING DATE: 2003-08-29  
; PRIOR APPLICATION NUMBER: US09/839,650  
; PRIOR FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: US09/748,786  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Humanized R. mulleri polynucleotide  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(720)  
; OTHER INFORMATION: Humanized DNA sequence  
US-10-652-703A-1

Query Match 11.8%; Score 126; DB 3; Length 720;  
Best Local Similarity 51.3%; Pred. No. 9.7e-31;  
Matches 325; Conservative 0; Mismatches 300; Indels 9; Gaps 1;  
QY 123 ATGAGATCAAGCTCGTATGGAAGTGTGTAAACGGGCACAAAGTTCTGGTTGAAAGGA 182  
| | | | | | | | | | | | | | | | | | | | | |  
Db 46 ATGAGCTCAAGGTGAACCTGGAGGGCATCGTGAACAAACACCGTGTTCACCATGGAGGGC 105  
| | | | | | | | | | | | | | | | | | | | | |  
QY 183 GATGGAAGGGAAGGCTTTTCAGCGGAACACAGACTATGACCTTACAGTCATAGAAGGC 242  
| | | | | | | | | | | | | | | | | | | | | |  
Db 106 TCGGCAAGGGCAACATCTCTTTCGCAACACAGCTGGTCAGATCCGCGTGACCAAGGC 165  
| | | | | | | | | | | | | | | | | | | | | |  
QY 243 GCACCATTCGCTTTTCGATATCTTGAACAACAGTATTCGATTAACGGCAACAGGGTA 302  
| | | | | | | | | | | | | | | | | | | | | |  
Db 166 GCCCCCTCGCTTCGCTTCGACATCGTGAGCCCGCTTCAGTACGGCAACCGCAC 225  
| | | | | | | | | | | | | | | | | | | | | |  
QY 303 TTCGCCAATACCCAGAACATAGCAGATTATTTCAACAGACGTTTCTCGAGGGTAC 362  
| | | | | | | | | | | | | | | | | | | | | |  
Db 226 TTCACCAAGTACCCCAACAGCATCAGCGACTACTTCATCCAGAGCTTCCCGCGGCTTC 285  
| | | | | | | | | | | | | | | | | | | | | |  
QY 363 TTCTGGGACGAGCATGACATACAGACAGGCGATTTCGATCGCCACAAACGACATA 422  
| | | | | | | | | | | | | | | | | | | | | |  
Db 286 ATGTAGAGCGCACCTTCGCTACGAGGACGGCGCTCGTGAGATCCGACGCGCATC 345  
| | | | | | | | | | | | | | | | | | | | | |  
QY 423 ACAATGATGGAAGGGCTCGACGACTTTTGGCTATATAAATTCGATTTGATGTTGAAC 482  
| | | | | | | | | | | | | | | | | | | | | |  
Db 346 AACCTGA-----TCAGGCAAGTTCGTGTACCGGTGGAGTACAAAGGGCAGCAAC 396  
| | | | | | | | | | | | | | | | | | | | | |  
QY 483 TTTCCTGCCAATGGTTCAGTTATGACAGGAAGACGCTGAATGGAGGCAATCCACTGAG 542  
| | | | | | | | | | | | | | | | | | | | | |  
Db 397 TTCCCGGACGAGCGGCCCTGTATGCAAGAGACATCTCGGCATCGAGCCAGCTTCGAG 456  
| | | | | | | | | | | | | | | | | | | | | |  
QY 543 ATAATGATGCCGTGATGAGGTGTGAAGGTGATTTAAATGCTGTGTTGTTGA 602  
| | | | | | | | | | | | | | | | | | | | | |  
Db 457 GCCATGTATCAACAAACCGGCTGTGTGGGCGAGGTGATCTGTGTGTAAGCTGAAC 516  
| | | | | | | | | | | | | | | | | | | | | |  
QY 603 GGAGGTGGCCATTACCGATGTGACTTCAAACTACTTACAAAGCTAAGAGGTTGTCCGG 662  
| | | | | | | | | | | | | | | | | | | | | |

Db 517 AGCGCAAGTACTACAGCTGCCACATGAAGACCTTGATGAAGAGCAAGGCGTGGTGAAG 576  
| | | | | | | | | | | | | | | | | | | | | |  
QY 663 TTGCCAGACTATCAGCTTTTGTGGACCATCGCATTTGAGATTGTGAGCCACGACAAGATTAC 722  
| | | | | | | | | | | | | | | | | | | | | |  
Db 577 GAGTTCCCTCTCTACCACTTTCATCCAGCACCGCTCGAAGAACCTTACGTGGAGGACGGC 636  
| | | | | | | | | | | | | | | | | | | | | |  
QY 723 AACAAAGTTAAGCTGCACGAGCATGCCGAAGCTC 756  
| | | | | | | | | | | | | | | | | | | | | |  
Db 637 GGCTTCGTGGAGCAGCAGCAGCCGCGCATCGCCC 670

Search completed: April 5, 2006, 22:28:00  
Job time : 237 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 21:02:02 ; Search time 1060 Seconds

(without alignments)  
8316.186 Million cell updates/sec

Title: US-10-757-356-17

Perfect score: 1066

Sequence: 1 attcgccctgggatttggg.....atcgagcgcaaaaaaaa 1066

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	1066	100.0	1066	8	US-10-757-356-17		Sequence 17, Appl
2	838.8	78.7	913	8	US-10-757-356-25		Sequence 25, Appl
3	665	62.4	1133	8	US-10-757-356-27		Sequence 27, Appl
4	628	59.9	684	6	US-10-314-827-21		Sequence 21, Appl
5	628	58.9	684	7	US-10-314-936-7		Sequence 7, Appl
6	628	58.9	684	10	US-11-021-014-7		Sequence 7, Appl
7	626.4	58.8	684	7	US-10-314-936-5		Sequence 5, Appl
8	626.4	58.8	684	10	US-11-021-014-5		Sequence 5, Appl
9	533.4	50.0	795	8	US-10-757-356-15		Sequence 15, Appl
10	517.2	48.5	860	5	US-10-244-779-1		Sequence 1, Appl
11	507.4	47.6	729	9	US-10-851-636-12		Sequence 12, Appl
12	507.2	47.6	678	9	US-10-851-636-4		Sequence 4, Appl
13	497.8	46.7	729	9	US-10-851-636-13		Sequence 13, Appl
14	497.4	46.7	678	9	US-10-851-636-5		Sequence 5, Appl
15	486.4	45.6	678	9	US-10-492-081B-2		Sequence 2, Appl
16	467.2	43.8	675	9	US-10-851-636-3		Sequence 3, Appl
17	467.2	43.8	726	6	US-10-851-636-11		Sequence 11, Appl
18	465	43.6	726	6	US-10-314-827-3		Sequence 3, Appl
19	461.8	43.3	726	6	US-10-314-827-5		Sequence 5, Appl
20	460.2	43.2	726	6	US-10-314-827-13		Sequence 13, Appl
21	458.6	43.0	726	6	US-10-314-827-11		Sequence 11, Appl
22	457	42.9	726	6	US-10-314-827-9		Sequence 9, Appl
23	455.4	42.7	726	6	US-10-314-827-7		Sequence 7, Appl

24	436	40.9	745	6	US-10-314-827-17	Sequence 17, Appl
25	436	40.9	746	6	US-10-314-827-15	Sequence 15, Appl
26	433	40.6	748	6	US-10-314-827-19	Sequence 19, Appl
27	432.8	40.6	684	7	US-10-664-341-68	Sequence 68, Appl
28	431.6	40.5	681	6	US-10-314-827-1	Sequence 1, Appl
29	419.6	39.4	801	3	US-09-967-772-7	Sequence 7, Appl
30	419.6	39.4	801	6	US-10-335-517-7	Sequence 7, Appl
31	419.6	39.4	801	6	US-10-334-288-7	Sequence 6, Appl
32	419.6	39.4	801	7	US-10-311-030-6	Sequence 6, Appl
33	419.6	39.4	801	8	US-10-656-029-23	Sequence 23, Appl
34	419.6	39.4	1116	5	US-10-006-922-3	Sequence 3, Appl
35	418.6	39.3	845	8	US-10-757-356-3	Sequence 3, Appl
36	409.2	38.4	741	9	US-10-851-636-14	Sequence 14, Appl
37	405.4	38.0	693	9	US-10-851-636-6	Sequence 6, Appl
38	380	35.6	725	9	US-10-851-636-39	Sequence 39, Appl
39	380	35.6	726	9	US-10-851-636-38	Sequence 38, Appl
40	380	35.6	726	9	US-10-851-636-43	Sequence 43, Appl
41	378.4	35.5	725	9	US-10-851-636-41	Sequence 41, Appl
42	378.4	35.5	726	9	US-10-851-636-40	Sequence 40, Appl
43	378.4	35.5	726	9	US-10-851-636-44	Sequence 44, Appl
44	375.2	35.2	726	9	US-10-851-636-42	Sequence 42, Appl
45	371.8	34.9	680	9	US-10-851-636-91	Sequence 91, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-757-356-17  
; Sequence 17, Application US/10757356  
; Publication No. US20050032085A1  
; GENERAL INFORMATION:  
; APPLICANT: Labae, Yulii Aleksandrovich  
; APPLICANT: Gurskaya, Nadezda Georgievna  
; APPLICANT: Yanushevich, Yuriy  
; APPLICANT: Fradkov, Arcady Fedorovich  
; APPLICANT: Lukyanov, Konstantin  
; APPLICANT: Lukyanov, Sergey  
; APPLICANT: Matz, Mikhail Vladimirovich  
; TITLE OF INVENTION: NOVEL CHROMOPHORES/FLUOROPHORES AND  
; FILE REFERENCE: CLON-090  
; CURRENT APPLICATION NUMBER: US/10/757,356  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: 60/332,980  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: PCT/US02/36499  
; PRIOR FILING DATE: 2002-11-12  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 1066  
; TYPE: DNA  
; ORGANISM: Montastraea cavernosa  
US-10-757-356-17

Query Match 100.0%; Score 1066; DB 8; Length 1066;  
Best Local Similarity 100.0%; Pred. No. 1e-310;  
Matches 1066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ATTCGCTCGTGTGTTGGAGAGAGAGATCGAGAGACACAGAGCTGTAGGTTGATA 60  
Db 1 ATTCGCTCGTGTGTTGGAGAGAGAGATCGAGAGACACAGAGCTGTAGGTTGATA 60  
Oy 61 TCTTACTTACCTCTACCATCATGCAAGTGTTCACAGGAAAGGTTGATTAACCCAG 120  
Db 61 TCTTACTTACCTCTACCATCATGCAAGTGTTCACAGGAAAGGTTGATTAACCCAG 120  
Oy 121 ACATGAGATGAAGCTCGGTATGGAAGTGTCTTAACGGGCACAGTTCGTGTTGAAG 180  
Db 121 ACATGAGATGAAGCTCGGTATGGAAGTGTCTTAACGGGCACAGTTCGTGTTGAAG 180  
Oy 181 GAGATGGAAGGGAAGCCCTTCGACGGAACACAGACTATGACCTTACGCTACATAGAAG 240

```

181 GAGATGGAAGGAGGAGCTTTGACGGAACACAGACTATGACCTTACAGTCTATAGAG 240
241 GCGCACCATTCCTTTCCTTACGATATCTTGACAAACAGTATTCGATTACGGCAACAGGG 300
241 GCGCACCATTCCTTTCCTTACGATATCTTGACAAACAGTATTCGATTACGGCAACAGGG 300
301 TATTGCGCAATATCCAGAACATAGCAGATATTTCAAGCAGACGTTTCCCTGAGGGT 360
301 TATTGCGCAATATCCAGAACATAGCAGATATTTCAAGCAGACGTTTCCCTGAGGGT 360
361 ACTTCTGGAACGAGCATGACATACGAGAACACGAGGCAATTTGATCGCCACAAACGACA 420
361 ACTTCTGGAACGAGCATGACATACGAGAACACGAGGCAATTTGATCGCCACAAACGACA 420
421 TAAACAATGATGGAAGGCTGACAGCTGTTTGGCTATATAAATTCGATTGATGGTGTGA 480
421 TAAACAATGATGGAAGGCTGACAGCTGTTTGGCTATATAAATTCGATTGATGGTGTGA 480
481 ACTTCTGGAACGAGCATGACATACGAGAACACGAGGCAATTTGATCGCCACAAACGACA 540
481 ACTTCTGGAACGAGCATGACATACGAGAACACGAGGCAATTTGATCGCCACAAACGACA 540
541 AGATAATGATGCGCGTATGAGTGTGAGGCTGATGTTTAAACATGGCTCTGTTGCTTG 600
541 AGATAATGATGCGCGTATGAGTGTGAGGCTGATGTTTAAACATGGCTCTGTTGCTTG 600
601 AAGGAGGTGGCCATTTACCGATGTGACTTCAAACTTCAAACTTCAAACTTCAAACTT 660
601 AAGGAGGTGGCCATTTACCGATGTGACTTCAAACTTCAAACTTCAAACTTCAAACTT 660
661 GGTTCGCAAGTATCACTTTGAGCACTATGAGTGTGAGGCAATTTGATCGCCACAAACGACA 720
661 GGTTCGCAAGTATCACTTTGAGCACTATGAGTGTGAGGCAATTTGATCGCCACAAACGACA 720
721 ACAACAAGTTAAGTGCACGAGCATGCGAAGTGTGCTGATGAGTGTGAGGCAATTTGATCG 780
721 ACAACAAGTTAAGTGCACGAGCATGCGAAGTGTGCTGATGAGTGTGAGGCAATTTGATCG 780
781 AGTAAAGCTTAAATGAAAGTCAAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
781 AGTAAAGCTTAAATGAAAGTCAAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
841 AATTGAGGCAATTTACTCGGAATTAGTATTTGATGATGATGATGATGATGATGATGATGAT 900
841 AATTGAGGCAATTTACTCGGAATTAGTATTTGATGATGATGATGATGATGATGATGATGAT 900
901 GGATTTGTTAGAGTACTCTAGAGTGTGATTTTGTGAAAGGAGATAGTTTCCAGTTT 960
901 GGATTTGTTAGAGTACTCTAGAGTGTGATTTTGTGAAAGGAGATAGTTTCCAGTTT 960
961 TCGCGGATTTACAGCATGGGATAGACTTTTAACTCAGTTTGTGCTCAAAATGCAAGTAAG 1020
961 TCGCGGATTTACAGCATGGGATAGACTTTTAACTCAGTTTGTGCTCAAAATGCAAGTAAG 1020
1021 AAAACTGTAGTGAATTAACCTGTTATCGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1066
1021 AAAACTGTAGTGAATTAACCTGTTATCGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1066

```

RESULT 2

```

US-10-757-356-25
; Sequence 25, Application US/10757356
; Publication No. US20050032085A1
; GENERAL INFORMATION:
; APPLICANT: Labas, Yulii Alekseevich
; APPLICANT: Gurskaya, Nadezda Georgievna
; APPLICANT: Yanushevich, Yuriy
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Matz, Mikhail Vladimirovich
; TITLE OF INVENTION: NOVEL CHROMOPHORES/FLUOROPHORES AND

```

```

; TITLE OF INVENTION: METHODS FOR USING THE SAME
; FILE REFERENCE: CLON-090
; CURRENT APPLICATION NUMBER: US/10/757,356
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: 60/332,980
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/US02/36499
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 913
; TYPE: DNA
; ORGANISM: Montastrea cavernosa
; US-10-757-356-25

Query Match 78.7%; Score 838.8; DB 8; Length 913;
Best Local Similarity 96.9%; Pred. No. 4e-242;
Matches 866; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

Qy 52 AGTTGATATCTTACTTACGTCTACCATCATGACAGTGTTCACAGGAAAGGGTGTGA 111
Db 9 AGGGTGATATCTTACTTACGTCTACCATCATGACAGTGTTCACAGGAAAGGGTGTGA 68
Qy 112 TTAACCCAGACATGAAGATGAAGCTGCTATGGAAGTGTCTGTAACCGGCACAAAGTTTCG 171
Db 69 TTAACCCAGACATGAAGATGAAGCTGCTATGGAAGTGTCTGTAACCGGCACAAAGTTTCG 128
Qy 172 TGGTTGAAGGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 231
Db 129 TGATTGAAGGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 188
Qy 232 TCATAGAGGCGCACCATTCGCTTTACGATATCTTACAGTATCTTACAGGAGTATTCGATTACG 291
Db 189 TCATAGAGGCGCACCATTCGCTTTACGATATCTTACAGTATCTTACAGGAGTATTCGATTACG 248
Qy 292 GCAACAGGGTATTCGCAAAATACCCAGAGACATAGCAGATATTTTCAAGCAGACGTTTC 351
Db 249 GCAACAGGGTATTCGCAAAATACCCAGAGACATAGCAGATATTTTCAAGCAGACGTTTC 308
Qy 352 CTGAGGGGTACTTCTGGGAAACGAGCATGACATACGAAAGACCGAGGAGGAGGAGGAGGAG 411
Db 309 CTGAGGGGTACTTCTGGGAAACGAGCATGACATACGAAAGACCGAGGAGGAGGAGGAGGAG 368
Qy 412 CAACAGCATACCAATGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 471
Db 369 CAACAGCATACCAATGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 428
Qy 472 ATGGTGTGAAGTATTCGCTCAAAATGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 531
Db 429 ATGGTGTGAAGTATTCGCTCAAAATGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 488
Qy 532 CATCCACTGAGATAATGATGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 591
Db 489 CATCCACTGAGATAATGATGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 548
Qy 592 TGTGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 651
Db 549 TGTGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 608
Qy 652 AGTTGTCGCGTTGCGCAGACTATCACATTTCTGGAACCATTCGATTTGAGATTTGTGAGCCACG 711
Db 609 AGTTGTCGCGTTGCGCAGACTATCACATTTCTGGAACCATTCGATTTGAGATTTGTGAGCCACG 668
Qy 712 ACAAGATTAACAAGGTTTAAAGTGTGACGAGCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 771
Db 669 ACAAGATTAACAAGGTTTAAAGTGTGACGAGCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 728
Qy 772 GGAAGGCAAGTAAAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 830
Db 729 GGAAGGCAAGTAAAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 788
Qy 831 TTTTGTGTTAAATTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 890

```

Db 789 TTTTGTGTTAAATTTCAAGGCAATTTACTCGGAATTTAGTATTGTAATTTTCGATTCAAGG 848  
 Qy 891 ATTGTTCCGGGATTTGTTAGAGCTAGCTCTAGAGTTGTTATTTTGTGAAAAA 944  
 Db 849 ATTGTTCCGGGACTTTGTTAGAGACCACTCTAGAGTTGTTATTTTGTGAAAAA 902

RESULT 3

US-10-757-356-27  
 ; Sequence 27, Application US/10757356  
 ; Publication No. US20050032085A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Labas, Yulii Alekseevich  
 ; APPLICANT: Gurskaya, Nadezda Georgievna  
 ; APPLICANT: Yanushevich, Yuriy  
 ; APPLICANT: Fradkov, Arcady Fedorovich  
 ; APPLICANT: Lukyanov, Konstantin  
 ; APPLICANT: Lukyanov, Sergey  
 ; APPLICANT: Matz, Mikhail Vladimirovich  
 ; TITLE OF INVENTION: NOVEL CHROMOPHORES/FLUOROPHORES AND  
 ; TITLE OF INVENTION: METHODS FOR USING THE SAME  
 ; FILE REFERENCE: CLON-090  
 ; CURRENT APPLICATION NUMBER: US/10/757,356  
 ; CURRENT FILING DATE: 2004-01-13  
 ; PRIOR APPLICATION NUMBER: 60/332,980  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: PCT/US02/36499  
 ; PRIOR FILING DATE: 2002-11-12  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 27  
 ; LENGTH: 1133  
 ; TYPE: DNA  
 ; ORGANISM: Montastraea annularis  
 US-10-757-356-27

Query Match 62.4%; Score 665; DB 8; Length 1133;  
 Best Local Similarity 80.4%; Pred. No. 1.5e-189;  
 Matches 855; Conservative 2; Mismatches 162; Indels 44; Gaps 5;

Qy 1 ATTCGCCCTGGTGTGATTTGGAGAGAGCAGATCGAGAACACAAAGAGCTGTA--AGGTGA 58  
 Db 99 ATTCACCTGGTGTGATTTGGAGAGAGCAGATCGAGAACACAAAGAGCTGTTATCGCTAA 158  
 Qy 59 TATCTTACTTACGCTTACCATCATGACAAAGTGTTCGACAGGAAAGGGTGTGATTAACC 118  
 Db 159 AATCTTACTTGGCTCTACCAACATGA-----GTATGATTAACC 197  
 Qy 119 AGACATGAAGATGAAGTCCGCTATGAGAGTCTGTAAACGGGCACAACTTCGTGTTGA 178  
 Db 198 AGAAATGAAGATCAAGATCGGTATGAGCGGTCTGTAAACGGGCACAACTTCGTGATTAC 257  
 Qy 179 AGGAGATGAAAGGGAAGCCCTTCGACGGAAACACAGACTATGACCTTACAGTCAATAGA 238  
 Db 258 AGGGAAGGAAGCGGGAGCCTTCGAGGGAACACAGACTATGAACCTGACAGTCAATAGA 317  
 Qy 239 AGGCGCACCATTCGCTTCGCTTACGATATCTTGAACACAGTATTCGATTCGATCGGCAACAG 298  
 Db 318 CGGCGGACCTTCGCTTCGCTTACGATATCTTGAACACAGTATTCGATTCGATTCGCGCAGC 377  
 Qy 299 GGTATTCGCAATATCCAGAGACATAGCAGATTTTTCAGACAGAGCTTCCTCGAGGG 358  
 Db 378 GGTATTCGCAATATCCAGAGACATCCGAGCTATTTCAAGAGCTTCCTTCGAGGG 437  
 Qy 359 GTACTTCTGGGAACGAAGCATGACATACGAAGACAGAGGCTTCGATCGCCACAAACGA 418  
 Db 438 GTTCTTCTGGNACGAGAGCATGACTTACGAAGCGGGGCACTTCGATCGCCACAAATGA 497  
 Qy 419 CATAAATATGATGGAAGGCGTCCAGCAGCTGTTTTCCTATATAAATTCGATTTGATGGTGT 478  
 Db 498 CATAAAAATG-----GAAGGCGAGCTCTTTCTCTATGAAATTCGATTTGATGGGT 548

Qy 479 GAACTTTCTCCCAATGGTCCAGTTATGTCAGAGGAAGACGCTGAAATGGAGCCATCCAC 538  
 Db 549 GAACTTTCTCCCAATAGTCCAGTTATGTCAGAGGAAGACCGTGAATGGAGCCATCCAC 608  
 Qy 539 TGAGATATGATGCGGTGATGAGTGTCTGAAGGGTGTGTTAAACATGCTCTCTTGTCT 598  
 Db 609 TGRGAAATGATGTCGCTGATGAGTGTCTTAAAGGGTGTCTTAAACATGCTCTCTTGTCT 668  
 Qy 599 TGRAGGAGTGGCCATTTACCGATGCTGAACTTCCAAACACTTCTACAAAGCTTAAGAAGTGT 658  
 Db 669 TGAAGGAGTGGCCATTTCCGATGCTGAACTTGAACAACTTCTACAAAGCTTAAGAAGTGT 728  
 Qy 659 CCGGTGTCAGACTATCACCTTTGTGGACCATCGCATTCGAGATTTGTGAGCCACGACAAAGA 718  
 Db 729 CCAGATCCAGACTATCACCTTTGTGAATCACCGACTTGAGATAACATGSCATGACGAGGA 788  
 Qy 719 TTACAAAGGTTAAGTGTGACGAGCATGCCGAGCTGTCTGATGCTCATGAGTGTCAAGGAAGGC 778  
 Db 789 TTACAACAATGTTAAGCTGTCTGAGCATGCAAGAGCTCATTTCTGGACTGCCAAGCAGGC 848  
 Qy 779 CAAGTAAAGGCTTAATGAAAGTCAAGACGACCAACGAGGAGAAACA-----AAGTA 829  
 Db 849 CAATAAAGGCTTGACGAAAGCCAAACGCGCAAGAGTACAGAAAGTATATATAATG 908  
 Qy 830 CTTTTTTGTTAAATTTGAAGGCATTTTACTCGGAATTTAGTATTTGATACTTTTCGATTCAAG 889  
 Db 909 TATATTTTCACTGAAAGGCATTTCCACTCGGAATTTAGTATTTGATACTTTTCAATTCAAG 968  
 Qy 890 GATTGTTCCGGGATTTGTTAGAGCTAGCTCTAGAGTGTGATTTTGTGAAAAAAGATAG 949  
 Db 969 GATTATTTTGGGATTTGCTAGCCACTAGCTTTATTTGTTAAATTAAGTTAAAGACG---G 1025  
 Qy 950 TTTCCAGTTTTTGGGGATTTACAGCTGGGATAGACTTTTTTAAACTCAGTTGTCGTCAA 1009  
 Db 1026 TTTAGCATTTTTTCGGTATTAACAACATAGGACAGACGCTTTAACCCAGATAGTGTCTAG 1085  
 Qy 1010 ATCAAGTAAGAAACTGTAGTGAATAAACTTTGTTATCGAA 1052  
 Db 1086 GTACAGTAAGAAACTTTGTTGAGATAGACTTTGTTAGTCGAA 1128

RESULT 4

US-10-314-827-21  
 ; Sequence 21, Application US/10314827  
 ; Publication No. US20030157643A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Almond, Brian D.  
 ; APPLICANT: Wood, Monika G.  
 ; APPLICANT: Wood, Keith V.  
 ; TITLE OF INVENTION: SYNTHETIC NUCLEIC ACIDS FROM AQUATIC SPECIES  
 ; FILE REFERENCE: 638,005  
 ; CURRENT APPLICATION NUMBER: US/10/314,827  
 ; CURRENT FILING DATE: 2002-12-09  
 ; PRIOR APPLICATION NUMBER: 09/645706  
 ; PRIOR FILING DATE: 2000-08-24  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 21  
 ; LENGTH: 684  
 ; TYPE: DNA  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: parent  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(684)  
 ; OTHER INFORMATION:  
 US-10-314-827-21

Query Match 58.9%; Score 628; DB 6; Length 684;  
 Best Local Similarity 94.9%; Pred. No. 1.7e-178;  
 Matches 649; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

102 AAGGGTGTGATTAACACGACATGAAGTGAAGTGTGATGAAGGCTGTAAACGGG 161  
 1 ATGAGTGTGATTAACACGACATGAAGTGAAGTGTGATGAAGGCTGTAAACGGG 60  
 162 CACAAAGTTCGTGTTGAAGGAGATGGAAGGAGGAGCTTTGACGGAACACAGACTATG 221  
 61 CACAAGTTCGTGATTTGAAGGAGATGGAAGGAGGAGCTTTGACGGAAGGAGACTATG 120  
 222 GACCTTACAGTTCATAGAGGAGGAGCTTTGCTTTCGCTTACGATATCTTTGACAAACAGTA 281  
 121 GACCTTACAGTTCATAGAGGAGGAGCTTTGCTTTCGCTTACGATATCTTTGACAAACAGTA 180  
 282 TTCGATTACGCGCAACAGGGTATTCCGCAAAATACCCAGAAACATACGAGACTATTTCAAG 341  
 181 TTCGATTACGCGCAACAGGGTATTCCGCAAAATACCCAGAAACATACGAGACTATTTCAAG 240  
 342 CAGACGTTTCTGAGGGGTACTTCTGGAAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 401  
 241 CAGACGTTTCTGAGGGGTACTTCTGGAAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 300  
 402 TGCATGCGCCACAAACGACATGAAGTGAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 461  
 301 TGCATGCGCCACAAACGACATGAAGTGAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 360  
 462 ATTCGATTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 521  
 361 ATTCGATTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 420  
 522 AATGAGGAGGAGTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 581  
 421 AATGAGGAGGAGTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 480  
 582 AACATGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 641  
 481 AACATGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 540  
 642 AAGCTAAGAGAGTGTGCGGTTGCCAGACTATCATTGTTGAGACCATCGCATGAGATT 701  
 541 AAGCTAAGAGAGTGTGCGGTTGCCAGACTATCATTGTTGAGACCATCGCATGAGATT 600  
 702 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGATGAGCTGATGAGCTGCGGAGCTGTCAT 761  
 601 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGATGAGCTGATGAGCTGCGGAGCTGTCAT 660  
 762 GGACTGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 785  
 661 GGAGTGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684

RESULT 5

US-10-314-936-7  
 ; Sequence 7, Application US/10314936  
 ; Publication No. US20040110225A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gibbs, Patrick D.L.  
 ; APPLICANT: Carter, Robert W.  
 ; APPLICANT: Schmale, Michael C.  
 ; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
 ; FILE REFERENCE: 638.004  
 ; CURRENT APPLICATION NUMBER: US/10/314,936  
 ; CURRENT FILING DATE: 2002-12-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 7  
 ; LENGTH: 684  
 ; TYPE: DNA  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: mutant green fluorescent protein  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(684)  
 ; OTHER INFORMATION:

US-10-314-936-7  
 Query Match 58.9%; Score 628; DB 7; Length 684;  
 Best Local Similarity 94.9%; Pred. No. 1.7e-178; Indels 0; Gaps 0;  
 Matches 649; Conservative 0; Mismatches 35;  
 QY 102 AAGGGTGTGATTAACACGACATGAAGTGAAGTGTGATGAAGGCTGTAAACGGG 161  
 Db 1 ATGAGTGTGATTAACACGACATGAAGTGAAGTGTGATGAAGGCTGTAAACGGG 60  
 QY 162 CACAAAGTTCGTGTTGAAGGAGATGGAAGGAGGAGCTTTGACGGAACACAGACTATG 221  
 Db 61 CACAAGTTCGTGATTTGAAGGAGATGGAAGGAGGAGCTTTGACGGAAGGAGACTATG 120  
 QY 222 GACCTTACAGTTCATAGAGGAGGAGCTTTGCTTTCGCTTACGATATCTTTGACAAACAGTA 281  
 Db 121 GACCTTACAGTTCATAGAGGAGGAGCTTTGCTTTCGCTTACGATATCTTTGACAAACAGTA 180  
 QY 282 TTCGATTACGCGCAACAGGGTATTCCGCAAAATACCCAGAAACATACGAGACTATTTCAAG 341  
 Db 181 TTCGATTACGCGCAACAGGGTATTCCGCAAAATACCCAGAAACATACGAGACTATTTCAAG 240  
 QY 342 CAGACGTTTCTGAGGGGTACTTCTGGAAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 401  
 Db 241 CAGACGTTTCTGAGGGGTACTTCTGGAAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 300  
 QY 402 TGCATGCGCCACAAACGACATGAAGTGAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 461  
 Db 301 TGCATGCGCCACAAACGACATGAAGTGAAGGAGGAGTATGATACGAGACCAAGGGGCAAT 360  
 QY 462 ATTCGATTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 521  
 Db 361 ATTCGATTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 420  
 QY 522 AATGAGGAGGAGTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 581  
 Db 421 AATGAGGAGGAGTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 480  
 QY 582 AACATGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 641  
 Db 481 AACATGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 540  
 QY 642 AAGCTAAGAGAGTGTGCGGTTGCCAGACTATCATTGTTGAGACCATCGCATGAGATT 701  
 Db 541 AAGCTAAGAGAGTGTGCGGTTGCCAGACTATCATTGTTGAGACCATCGCATGAGATT 600  
 QY 702 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGATGAGCTGATGAGCTGCGGAGCTGTCAT 761  
 Db 601 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGATGAGCTGATGAGCTGCGGAGCTGTCAT 660  
 QY 762 GGACTGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 785  
 Db 661 GGAGTGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684

RESULT 6

US-11-021-014-7  
 ; Sequence 7, Application US/11021014  
 ; Publication No. US20050100954A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gibbs, Patrick D.L.  
 ; APPLICANT: Carter, Robert W.  
 ; APPLICANT: Schmale, Michael C.  
 ; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
 ; FILE REFERENCE: 638.004  
 ; CURRENT APPLICATION NUMBER: US/11/021,014  
 ; CURRENT FILING DATE: 2004-12-23  
 ; PRIOR APPLICATION NUMBER: US/10/314,936  
 ; PRIOR FILING DATE: 2002-12-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 7  
 ; LENGTH: 684

; TYPE: DNA  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: mutant green fluorescent protein  
 ;  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(684)  
 ; OTHER INFORMATION:  
 US-11-021-014-7

Query Match 58.8%; Score 628; DB 10; Length 684;  
 Best Local Similarity 94.9%; Pred. No. 1.7e-178;  
 Matches 649; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
 QY 102 AAGGGTGTGATTAAACACAGACATGAAGTCAAGCTGCTGATGAAGGTCTGTAACGGG 161  
 Db |  
 QY 1 ATGAGTGTGATTAACACAGACATGAAGTCAAGCTGCTGATGAAGGTCTGTAACGGG 60  
 Db |  
 QY 162 CACAAGTTCGTGTTGAAGGAGATGGAAGGAAAGCCCTTTCGACGGAAACACAGACTATG 221  
 Db |  
 QY 61 CACAAGTTCGTGTTGAAGGAGATGGAAGGAAAGCCCTTTCGACGGAAACACAGACTATG 120  
 Db |  
 QY 222 GACCTTACAGTTCATAGAACGCGCACCTTTCGCTTTCGCTTACGATATCTTGACACAGTA 281  
 Db |  
 QY 121 GACCTTACAGTTCATAGAACGCGCACCTTTCGCTTTCGCTTACGATATCTTGACACAGTA 180  
 Db |  
 QY 282 TTCGATTACGGCAACAGGGGTATTTCGCCAATACCCAGAAAGACATAGCAGATTATTTCAG 341  
 Db |  
 QY 181 TTCGATTACGGCAACAGGGGTATTTCGCCAATACCCAGAAAGACATAGCAGATTATTTCAG 240  
 Db |  
 QY 342 CAGACGTTTCCGAGGGGTACTCTGGGAAACGAAGCATGACATACGAACACAGGGGCATT 401  
 Db |  
 QY 241 CAGACGTTTCCGAGGGGTACTCTGGGAAACGAAGCATGACATACGAACACAGGGGCATT 300  
 Db |  
 QY 402 TGCATCGCCACAAACGACATACATGATGGAAGGCGTCGACGACTGTTTGCCTTATAA 461  
 Db |  
 QY 301 TGCATCGCCACAAACGACATACATGATGGAAGGCGTCGACGACTGTTTGCCTTATAA 360  
 Db |  
 QY 462 ATTCGATTGTGTTGATGTTGAACTTCTCGCAATGTCGAGTTATCAGAGGAAGACGCTG 521  
 Db |  
 QY 361 ATTCGATTGTGTTGATGTTGAACTTCTCGCAATGTCGAGTTATCAGAGGAAGACGCTA 420  
 Db |  
 QY 522 AAATGGAGCCATCCACTGAGATAATGTATGCGCGTGTGAGTGTGCTGAAGGGTGATGT 581  
 Db |  
 QY 421 AAATGGAGCCATCCACTGAGATAATGTATGCGCGTGTGAGTGTGCTGAAGGGTGATGT 480  
 Db |  
 QY 582 AACATGGCTCTGTTGTTGAAGGAGTGGCCATTACCGATGTGACTTCAAACTACTTAC 641  
 Db |  
 QY 481 AACATGGCTCTGTTGTTGAAGGAGTGGCCATTACCGATGTGACTTCAAACTACTTAC 540  
 Db |  
 QY 642 AAAGCTAAGAGGTTGTCGGTTCGAGACTATCACTTGTGACCATCGCATTGAGATT 701  
 Db |  
 QY 541 AAAGCTAAGAGGTTGTCGGTTCGAGACTATCACTTGTGACCATCGCATTGAGATT 600  
 Db |  
 QY 702 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGCGACGAGCATGCCGAAGCTCGTCAT 761  
 Db |  
 QY 601 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGCGACGAGCATGCCGAAGCTCGTCAT 660  
 Db |  
 QY 762 GGACTGTCAAGGAAGGCCAAGTAA 785  
 Db |  
 QY 661 GGGCTGCCGAGGCGGCCAAGTAA 684

RESULT 7  
 US-10-314-936-5  
 ; Sequence 5, Application US/10314936  
 ; Publication No. US20040110225A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gibbs, Patrick D.L.  
 ; APPLICANT: Carter, Robert W.  
 ; APPLICANT: Schmale, Michael C.  
 ; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
 ; FILE REFERENCE: 638.004

; CURRENT APPLICATION NUMBER: US/10/314,936  
 ; CURRENT FILING DATE: 2002-12-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 684  
 ; TYPE: DNA  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: mutant green fluorescent protein  
 ;  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(684)  
 ; OTHER INFORMATION:  
 US-10-314-936-5

Query Match 58.8%; Score 626.4; DB 7; Length 684;  
 Best Local Similarity 94.7%; Pred. No. 5.1e-178;  
 Matches 648; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
 QY 102 AAGGGTGTGATTAAACACAGACATGAAGTCAAGCTGCTGATGAAGGTCTGTAACGGG 161  
 Db |  
 QY 1 ATGAGTGTGATTAACACAGACATGAAGTCAAGCTGCTGATGAAGGTCTGTAACGGG 60  
 Db |  
 QY 162 CACAAGTTCGTGTTGAAGGAGATGGAAGGAAAGCCCTTTCGACGGAAACACAGACTATG 221  
 Db |  
 QY 61 CACAAGTTCGTGTTGAAGGAGATGGAAGGAAAGCCCTTTCGACGGAAACACAGACTATG 120  
 Db |  
 QY 222 GACCTTACAGTTCATAGAACGCGCACCTTTCGCTTTCGCTTACGATATCTTGACACAGTA 281  
 Db |  
 QY 121 GACCTTACAGTTCATAGAACGCGCACCTTTCGCTTTCGCTTACGATATCTTGACACAGTA 180  
 Db |  
 QY 282 TTCGATTACGGCAACAGGGGTATTTCGCCAATACCCAGAAAGACATAGCAGATTATTTCAG 341  
 Db |  
 QY 181 TTCGATTACGGCAACAGGGGTATTTCGCCAATACCCAGAAAGACATAGCAGATTATTTCAG 240  
 Db |  
 QY 342 CAGACGTTTCCGAGGGGTACTCTGGGAAACGAAGCATGACATACGAACACAGGGGCATT 401  
 Db |  
 QY 241 CAGACGTTTCCGAGGGGTACTCTGGGAAACGAAGCATGACATACGAACACAGGGGCATT 300  
 Db |  
 QY 402 TGCATCGCCACAAACGACATACATGATGGAAGGCGTCGACGACTGTTTGCCTTATAA 461  
 Db |  
 QY 301 TGCATCGCCACAAACGACATACATGATGGAAGGCGTCGACGACTGTTTGCCTTATAA 360  
 Db |  
 QY 462 ATTCGATTGTGTTGATGTTGAACTTCTCGCAATGTCGAGTTATCAGAGGAAGACGCTG 521  
 Db |  
 QY 361 ATTCGATTGTGTTGATGTTGAACTTCTCGCAATGTCGAGTTATCAGAGGAAGACGCTA 420  
 Db |  
 QY 522 AAATGGAGCCATCCACTGAGATAATGTATGCGCGTGTGAGTGTGCTGAAGGGTGATGT 581  
 Db |  
 QY 421 AAATGGAGCCATCCACTGAGATAATGTATGCGCGTGTGAGTGTGCTGAAGGGTGATGT 480  
 Db |  
 QY 582 AACATGGCTCTGTTGTTGAAGGAGTGGCCATTACCGATGTGACTTCAAACTACTTAC 641  
 Db |  
 QY 481 AACATGGCTCTGTTGTTGAAGGAGTGGCCATTACCGATGTGACTTCAAACTACTTAC 540  
 Db |  
 QY 642 AAAGCTAAGAGGTTGTCGGTTCGAGACTATCACTTGTGACCATCGCATTGAGATT 701  
 Db |  
 QY 541 AAAGCTAAGAGGTTGTCGGTTCGAGACTATCACTTGTGACCATCGCATTGAGATT 600  
 Db |  
 QY 702 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGCGACGAGCATGCCGAAGCTCGTCAT 761  
 Db |  
 QY 601 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGCGACGAGCATGCCGAAGCTCGTCAT 660  
 Db |  
 QY 762 GGACTGTCAAGGAAGGCCAAGTAA 785  
 Db |  
 QY 661 GGGCTGCCGAGGCGGCCAAGTAA 684

RESULT 8  
 US-11-021-014-5  
 ; Sequence 5, Application US/11021014  
 ; Publication No. US20050100954A1

GENERAL INFORMATION:  
; APPLICANT: Gibbs, Patrick D.L.  
; APPLICANT: Carter, Robert W.  
; APPLICANT: Schmale, Michael C.  
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
; FILE REFERENCE: 638.004  
; CURRENT APPLICATION NUMBER: US/11/021.014  
; CURRENT FILING DATE: 2004-12-23  
; PRIOR APPLICATION NUMBER: US/10/314,936  
; PRIOR FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: mutant green fluorescent protein  
; NAME/KEY: CDS  
; LOCATION: (1)..(684)  
; OTHER INFORMATION:  
US-11-021-014-5

Query Match 58.8%; Score 626.4; DB 10; Length 684;  
Best Local Similarity 94.7%; Pred. No. 5.1e-178;  
Matches 648; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
QY 102 AAGGTGTGATTAACGACGATGAAGTGGTATGAGGTGCTGTAAACGGG 161  
DB 1 ATGAGTGTGATAAACGACGATGAAGTGGTATGAGGTGCTGTAAACGGG 60  
QY 162 CACAAGTTCGTGTTGAAGGAGATGGAAGGAGGCTTTGACGGAACACAGACTATG 221  
DB 61 CACAAGTTCGTGTTGAAGGAGATGGAAGGAGGCTTTGACGGAACACAGACTATG 120  
QY 222 GACCTTACAGTCTAGAGGCGCACCTATGCTTTGCTTTACGATATCTTTGACAAACAGTA 281  
DB 121 GACCTTACAGTCTAGAGGCGCACCTTTGCTTTGCTTTACGATATCTTTGACAAACAGTA 180  
QY 282 TTGATTTACGCAACAGGCTATTCGCAATATCCCAAGACATAGCAGATTTTCAAG 341  
DB 181 TTGATTTACGCAACAGGCTATTCGCAATATCCCAAGACATAGCAGATTTTCAAG 240  
QY 342 CAGAAGTTCCTGAGGCTACTCTGGAACGAGAGTATGATACGAGGACCGGCTAT 401  
DB 241 CAGAAGTTCCTGAGGCTACTCTGGAACGAGAGTATGATACGAGGACCGGCTAT 300  
QY 402 TGCATCGCCACAAACGACATAAATGATGGAAGGCGTGCAGCAGTCTTTTGGCTATAAA 461  
DB 301 TGCATCGCCACAAACGACATAAATGATGGAAGGCGTGCAGCAGTCTTTTGGCTATAAA 360  
QY 462 ATTGATTTGATGTTGAACTTCTTCGCAATGCTTATGAGGCTATGAGAGAGAGCGCTG 521  
DB 361 ATTGATTTGATGTTGAACTTCTTCGCAATGCTTATGAGGCTATGAGAGAGAGCGCTA 420  
QY 522 AATGAGGAGCATCTCACTGAGATTAATGATGCGGTGATGAGTGTGGAAGGCTGATGTT 581  
DB 421 AATGAGGAGCATCTCACTGAGATTAATGATGCGGTGATGAGTGTGGAAGGCTGATGTT 480  
QY 582 AACATGCTCTGTTGTTGAGGAGGTGGCCATTTACCGATGTGACTTTCAAACTACTTAC 641  
DB 481 AACATGCTCTGTTGTTGAGGAGGTGGCCATTTACCGATGTGACTTTCAAACTACTTAC 540  
QY 642 AAGCTAAGAGGTTGTCGGTTCGCAAGATATCACTTTTGGACCACTCGATTTGAGATT 701  
DB 541 AAGCTAAGAGGTTGTCGGTTCGCAAGATATCACTTTTGGACCACTCGATTTGAGATT 600  
QY 702 GTGAGCCACGCAAGATTACCAAGGTTAAGCTGACGAGCATGCGGAGCTCGTCAT 761  
DB 601 GTGAGCCACGCAAGATTACCAAGGTTAAGCTGATGAGCATGCGGAGCTCGTCAT 660  
QY 762 GGACTGTCAAGGAAGGCGCAAGTAA 785

Db 661 GGCTGCCGAGGCGAGCCCAAGTAA 684  
RESULT 9  
US-10-757-356-15  
; Sequence 15, Application US/10757356  
; Publication No. US20050032085A1  
; GENERAL INFORMATION:  
; APPLICANT: Labas, Yulii Aleksandrovich  
; APPLICANT: Gurskaya, Nadezda Georgievna  
; APPLICANT: Yanushevich, Yuriy  
; APPLICANT: Pradkov, Arcady Fedorovich  
; APPLICANT: Lukyanov, Konstantin  
; APPLICANT: Lukyanov, Sergey  
; APPLICANT: Matz, Mikhail Vladimirovich  
; TITLE OF INVENTION: NOVEL CHROMOPHORES/FLUOROPHORES AND  
; FILE REFERENCE: CLON-090  
; CURRENT APPLICATION NUMBER: US/10/757,356  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: 60/332,980  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: PCT/US02/36499  
; PRIOR FILING DATE: 2002-11-12  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 795  
; TYPE: DNA  
; ORGANISM: Montastraea cavernosa  
US-10-757-356-15  
Query Match 50.0%; Score 533.4; DB 8; Length 795;  
Best Local Similarity 83.2%; Pred. No. 6.9e-150;  
Matches 680; Conservative 0; Mismatches 101; Indels 36; Gaps 5;  
QY 1 ATTCGCCCTGGTGTGATTTGGAAGAGAGCAGATCGAGAACCAACAGAGCTG--TAAGGTTGA 58  
DB 9 ATTCGCCCTGGTGTGATTTGGAAGAGAGCAGACCCAGAGAACCAACAGAGCTGTATTAAGGCTGA 68  
QY 59 TATCTTAC--TTAGCTTACCATCATGCAAGTGTGTCACAGGNAAGGGTGTGATTAAC 117  
DB 69 TATCTTAC--TTAGCTTACCATCATGCAAGTGTGTCACAGGNAAGGGTGTGATTAAC 107  
QY 118 CAGACATGAAGATGAAGCTGCGTATGGAAGGCTGCTAAACGGGCAACAGTTCGTGTTG 177  
DB 108 CAGTCAATGAAGATGAAGCTGCGTATGGAAGGCTGCTAAACGGGCAACAGTTCGTGTTG 167  
QY 178 AAGGAGATGGAAGGAGGAGGAGGCTTTTCGACGGAACACAGACTATGGACCTTACAGTCAATG 237  
DB 168 TTGGAAGAGGAGGAGGAGGAGGCTTTTCGACGGAACACAGACTATGGACCTTACAGTCAATG 227  
QY 238 AAGGCGCACCATTCGCTTTCGCTTACGATATCTTGAACAAGTATTCGATTCGATTCGCGCAACA 297  
DB 228 AAGGCGCACCTCTGCTTTCGCTTACGATATCTTGAACAAGTATTCGATTCGATTCGCGCAACA 287  
QY 298 GGTATTTCCGCAAAATACCCAGAGACATAGCAGATTTTCAAGCAGAGAGCTTTCTCTGAGG 357  
DB 288 GGTATTTCCGCAAAATACCCAGAGACATAGCAGATTTTCAAGCAGAGAGCTTTCTCTGAGG 347  
QY 358 GGTATTTCCGCAAAATACCCAGAGACATAGCAGATTTTCAAGCAGAGAGCTTTCTCTGAGG 417  
DB 348 AGTATTTCCGCAAAATACCCAGAGACATAGCAGATTTTCAAGCAGAGAGCTTTCTCTGAGG 407  
QY 418 ACATAACAATGATGGAAGGCTGACGACTGTTTTCCTTATATAAATTCGATTTGATGGTG 477  
DB 408 AGATAACAATG-----GAAGGCGACTGTTTTTCAATAAAGTTCGATTTGATGGTG 458  
QY 478 TGAACCTTTCTGCCAATGCTCCAGTTCAGAGGAGAGAGCTGGAATGAGGAGGAGCTTCCA 537  
DB 459 TGAACCTTTCTGCCAATGCTCCAGTTCAGAGGAGAGAGCTGGAATGAGGAGGAGCTTCCA 518

QY 538 CTGAGATAATGTATGCGCGTGTGGAGTGTGTAAGGGTGTATGATTAACATGGCTCTGTTC 597  
DB 519 CTGAAAAATGTATGCGGTGTGGAGTGTGACGGGTGTATCAACATGGCTCTGTTC 578  
QY 598 TTGAGGAGGTGCGCAATACCGATGTGACTTCAAACTACTTACAAGCT---AAGAAG 654  
DB 579 TTGAAGGAGGTGCGCAATACCGATGTGACTTCAAACTACTTACAAGCTAAGAAGG 638  
QY 655 TTGTCGGTGGCAGACTACTCTTTGTGGACCATCGCTTGAAGTGTGAGCCAGCA 714  
DB 639 GTGTCAAGTTACAGATTATCACTTTGAGGATCACTCCATTGAGATTTGCGCCATGCA 698  
QY 715 AAGATTACAAAGGTTAAGCTGSCACGAGATGCGGAAGCTGTGATGACTGTCAAGGA 774  
DB 699 AAGATACACTGAGGTTAAGCTGTATGAGCATGCCGAAGCTCATTTCTGGCTGCCGAGG 758  
QY 775 AGCCCAAGTAAAGGCTTAATGAAAGTCAAGCAGCA 811  
DB 759 TGGCAAGTAAAGGCTTAACGAAAGCCAGACCA 795

RESULT 10  
US-10-244-779-1  
; Sequence 1, Application US/10244779  
; Publication No. US20030106078A1  
; GENERAL INFORMATION:  
; APPLICANT: Falkowski, Paul  
; APPLICANT: Sun, Yi  
; APPLICANT: Gorbunov, Maxim  
; APPLICANT: Wyman, Kevin  
; APPLICANT: Chen, Yi-Bu  
; TITLE OF INVENTION: mCPP Encoding Nucleic Acids,  
; Polypeptides, Antibodies and Methods of Use Thereof  
; FILE REFERENCE: Rut 00-002305  
; CURRENT APPLICATION NUMBER: US/10/244,779  
; CURRENT FILING DATE: 2002-12-19  
; PRIOR APPLICATION NUMBER: 60/322,189  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 860  
; TYPE: DNA  
; ORGANISM: Montastrea cavernosa  
US-10-244-779-1

Query Match 48.5%; Score 517.2; DB 5; Length 860;  
Best Local Similarity 80.5%; Pred. No. 5.7e-145;  
Matches 668; Conservative 0; Mismatches 128; Indels 34; Gaps 4;  
QY 1 ATTGCGCCCTGGTGATTGGGAAGAGCAGATCGAGAACCAAGAGCTGTAGGTTGATA 60  
DB 54 ATTCAACCTGGTGATTGGGAAGCGAATCTAGAAGAACAGCGCTGTAGAGCTGATA 113  
QY 61 TCTTAC-TTACGCTTACCATCATGACAAAGTGTGCACAGGAAAGGTTGATTAAACCA 119  
DB 114 TCTTACTTACGCTTACCATCATG-----GTGTGATTAAATCA 152  
QY 120 GACATGAAGATGAAGTGTGATGGAAGTGTGTAACCGGCAAGTTCGTGTGAA 179  
DB 153 GTCATGMAATCAAGTGTGATGGAAGCGCAATGTAAGCGGCAAGTTCATGATTACA 212  
QY 180 GGAGATGGAAGGAGGAGCTTTCGACGAAACACAGACTATGGACCTTACATCATAGNA 239  
DB 213 GGAGAGGTGGAAGGCAAGCTTTCGAGGGAACACACACTATAATCTTAAAGTCAAGAA 272  
QY 240 GCGGACCACTTCCCTTTCCTTACGATATCTTGAACAAAGTATTCATTACCGCAACAGG 299  
DB 273 GCGGACCTCTGCTTTCCTTACGATATCTTGAACAAAGTATTCATTACCGCAACAGG 332  
QY 300 GTATTTCGCAAAATACCCAGAAGACATAGCAGATTATTTCAAGCAGACGTTTCTGAGGG 359  
DB 333 GTATTTCACCAATATCCCAAAAGACATACAGACTATTTTCAAGCAGTCTGTTTCTGAGGG 392

QY 360 TACTTCTGGGAACGAAGCATGACATACGAAGACCAAGGCGCATTTGTCATCGCCACAACGAC 419  
DB 393 TATTCTCTGGGAAGGAAGCATGACTTTGGAAGACCAAGGCGTTTGCACCGTCACAGCGAC 452  
QY 420 ATAAACAATGATGAAGAGCGGTGCAAGCTGTGTTTGTCTATAAAATTCGATTTGATGTTG 479  
DB 453 ATAAAGTTG-----GAAGGCGACTGTGTTTTTCTACGAAATTCGATTTTATGTTG 503  
QY 480 AACTTTCCTGCATGTTCCAGTTATGAGGAGGAGACGCTGAAATGGAGGCAATCCACT 539  
DB 504 AACTTTCCTGCATGTTCCAGTTATGAGGAGGAGACGCTGAAATGGAGGCAATCCACT 563  
QY 540 GAGATAATGTATGCGCGTGTGAGTGTGAGGCGTGTGTTAAATCATGCTCTGTTGCTT 599  
DB 564 GAGATAATGTATGCGCGTGTGAGTGTGAGGCGTGTGTTAAATCATGCTCTGTTGCTT 623  
QY 600 GAAGAGGTGGCCATTACCGATGTGACTTCAAACTACTTAAAGAGTAAAGAGTTGTC 659  
DB 624 GAAGGGATAAACATCACCGATGTAATCTCAAGAGTACTTACGGGCAAAAGAGGTTGTC 683  
QY 660 CGTTGCCAGACTATCACTTTTGTGACCATGCAATGAGATTGTGAGCCACGACCAAGAT 719  
DB 684 GTGTTGCCAGATATCACTTTTGTGACCATGCAATGAGATTGTGAGCCACGACCAAGAT 743  
QY 720 TACAACAAGGTTAAGCTGCACGAGCATGCCGAAGCTCG---TCATGGAGCTGCAAGGAAG 776  
DB 744 TACAACAAGGTTAAGCTGCACGAGCATGCCGAAGCTCG---TCATGGAGCTGCAAGGAAG 803  
QY 777 GCCAAGTAAAGGCTTAATGAAAGTCAAGACGACCAAGAGGAGAAACAAA 826  
DB 804 GCCAAGTAAAGGCTTAATGAAAGGCCCCAACGACCAAGAGGAGAAACAAA 853  
RESULT 11  
US-10-851-636-12  
; Sequence 12, Application US/10851636  
; Publication No. US20050048609A1  
; GENERAL INFORMATION:  
; APPLICANT: Matz, Mikhail  
; APPLICANT: Kelmanson, Ilya  
; APPLICANT: Salih, Anya  
; APPLICANT: Meleshkevitch, Ella  
; TITLE OF INVENTION: Novel Fluorescent and Colored Proteins, and Polynucleotides that  
; TITLE OF INVENTION: Encode These Proteins  
; FILE REFERENCE: UP-364XCI  
; CURRENT APPLICATION NUMBER: US/10/851,636  
; CURRENT FILING DATE: 2004-05-20  
; PRIOR APPLICATION NUMBER: 60/472,196  
; PRIOR FILING DATE: 2003-05-20  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12  
; LENGTH: 729  
; TYPE: DNA  
; ORGANISM: Montastrea cavernosa  
US-10-851-636-12

Query Match 47.6%; Score 507.4; DB 9; Length 729;  
Best Local Similarity 85.8%; Pred. No. 4.7e-142;  
Matches 591; Conservative 0; Mismatches 86; Indels 12; Gaps 2;  
QY 102 AAGGTTGTGATTAAACCAAGATGAGATGAGCTGCTGATGGAAGTGTGTAACCGG 161  
DB 25 ATGAGTGTGATTAAACCAAGATGAGATGAGCTGCTGATGGAAGTGTGTAACCGG 84  
QY 162 CACAAGTTTCGTGTTGAAGGAGATGGAAGGAGGAGCCCTTCGACGGAACACAGACTATG 221  
DB 85 CACAACCTTCGTGATTGAAGGAGAGGAGGAGGAGCCCTTCGAGGGAACACAGACTATA 144  
QY 222 GACCTTACAGTCATAGAGGCGCACCATTTGCTTTGCTTTACGATATCTTGACAAACGTA 281  
DB 145 AACCTTACAGTCAAGAGGAGGAGCCCTCTGCTTTGCTTTACGATATCTTGACAGCAGCA 204





QY 222 GACCTTACAGTCATAGAGGCGCACATTTGCTTTTTCGCTTACGATATCTTGACACAGTA 281  
 DB 145 AACCTTACAGTCATAGAGGCGCACATCTCCCTTTTTCGCTTACGATATCTTGACACAGTA 204  
 QY 282 TTCGATTACGGCAACAGGGGTATTCGCCAAATACCCAGAGACATAGCAGATTTATTTCAAG 341  
 DB 205 TTCAGATAGGCAACAGGGGTATTCACCAATATCCAGACGATATACCAGATTTATTTCAAG 264  
 QY 342 CAGAGCTTTCCTGAGGGGTACTCTGCGGAGCGGAGTATGAGGAGTATGAGGAGTATTTCAAG 401  
 DB 265 CAGAGCTTTCCTGAGGGGTATTCGCGGAGCGGAGTATGAGGAGTATTTCAAG 324  
 QY 402 TGCATCGCCACAAACGACATCAATGAGGAGGCGTCCAGACGCTGTTTTCGCTTATATAA 461  
 DB 325 TGCAGCGCCACAAACGACATCAATGAGGAGGCGTCCAGACGCTGTTTTCGCTTATATAA 375  
 QY 462 ATTCGATTTCGAGGTGAGTATTCGCCAAATGCTCCAGATTTATGAGGAGGAGCGCTG 521  
 DB 376 ATTCGATTTCGAGGTGAGTATTCGCCAAATGCTCCAGATTTATGAGGAGGAGCGCTG 435  
 QY 522 AAATGGGAGCCATCCATGAGATATGATGCGGTGATGAGTGTGAGTGTGAGTGTGAGTGT 581  
 DB 436 AAATGGGAGCCATCCATGAGATATGATGCGGTGATGAGTGTGAGTGTGAGTGTGAGTGT 495  
 QY 582 AACATGGCTCTGCTGAGGAGGCGTCCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 641  
 DB 496 AACATGGCTCTGCTGAGGAGGCGTCCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 555  
 QY 642 AAAGCTTAAGAGGTGCTCGGTGTCAGACTATCACTTTGTCGACCATCGCATTCGATTCAGATT 701  
 DB 556 AAAGCTTAAGAGGTGCTCGGTGTCAGACTATCACTTTGTCGACCATCGCATTCGATTCAGATT 615  
 QY 702 GTGAGCCACGACAAAGATTACAAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAG 758  
 DB 616 TTGAGCCATGACAAAGATTACAAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAG 675  
 QY 759 CATGAGTGTCAAGGAGGCGCAAGTAAAG 787  
 DB 676 TCTATGCTGCGGAGTAAGGCGCAAGTAAAG 704

RESULT 14  
 US-10-851-636-5  
 ; Sequence 5, Application US/10851636  
 ; Publication NO. US20050048609A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Matz, Mikhail  
 ; APPLICANT: Kelmanson, Ilya  
 ; APPLICANT: Salih, Anya  
 ; APPLICANT: Meleshkevitch, Ella  
 ; TITLE OF INVENTION: Novel Fluorescent and Colored Proteins, and Polynucleotides that  
 ; TITLE OF INVENTION: Encode These Proteins  
 ; FILE REFERENCE: UF-364XC1  
 ; CURRENT APPLICATION NUMBER: US/10/851,636  
 ; CURRENT FILING DATE: 2004-05-20  
 ; PRIOR APPLICATION NUMBER: 60/472,196  
 ; PRIOR FILING DATE: 2003-05-20  
 ; NUMBER OF SEQ ID NOS: 96  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 5  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: Montastraea cavernosa  
 US-10-851-636-5

Query Match 46.7%; Score 497.4; DB 9; Length 678;  
 Best Local Similarity 85.0%; Pred. No. 4.8e-139;  
 Matches 594; Conservative 0; Mismatches 91; Indels 12; Gaps 2;  
 QY 102 AAGGGTGTGATTAAACACAGATGAAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 161  
 DB 1 ATGAGTGTGATTAAACACAGATGAAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 60

QY 162 CACAAAGTTTCGTTGTTGAAGGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 221  
 DB 61 CACAAAGTTTCGTTGTTGAAGGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 QY 222 GACCTTACAGTCATAGAGGCGCACATTTGCTTTTTCGCTTACGATATCTTGACACAGTA 281  
 DB 121 AACCTTACAGTCATAGAGGCGCACATCTCCCTTTTTCGCTTACGATATCTTGACACAGTA 180  
 QY 282 TTCGATTACGGCAACAGGGGTATTCGCCAAATACCCAGAGACATAGCAGATTTATTTCAAG 341  
 DB 181 TTCGATTACGGCAACAGGGGTATTCGCCAAATACCCAGAGATATACCAGATTTATTTCAAG 240  
 QY 342 CAGAGCTTTCCTGAGGGGTACTCTGCGGAGCGGAGTATGAGGAGTATTTATGAGGAGCGGATTT 401  
 DB 241 CAGAGCTTTCCTGAGGGGTACTCTGCGGAGCGGAGTATGAGGAGTATTTATGAGGAGCGGATTT 300  
 QY 402 TGCATCGCCACAAACGACATCAATGAGGAGGCGTCCAGACGCTGTTTTCGCTTATATAA 461  
 DB 301 TGCAGCGCCACAAACGACATCAATGAGGAGGCGTCCAGACGCTGTTTTCGCTTATATAA 351  
 QY 462 ATTCGATTTCGAGGTGAGTATTCGCCAAATGCTCCAGATTTATGAGGAGGAGCGCTG 521  
 DB 352 ATTCGATTTCGAGGTGAGTATTCGCCAAATGCTCCAGATTTATGAGGAGGAGCGCTG 411  
 QY 522 AAATGGGAGCCATCCATGAGATATGATGCGGTGATGAGTGTGAGTGTGAGTGTGAGTGT 581  
 DB 412 AAATGGGAGCCATCCATGAGATATGATGCGGTGATGAGTGTGAGTGTGAGTGTGAGTGT 471  
 QY 582 AACATGGCTCTGCTGAGGAGGCGTCCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 641  
 DB 472 AACATGGCTCTGCTGAGGAGGCGTCCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 531  
 QY 642 AAAGCTTAAGAGGTGCTCGGTGTCAGACTATCACTTTGTCGACCATCGCATTCGATTCAGATT 701  
 DB 532 AAAGCTTAAGAGGTGCTCGGTGTCAGACTATCACTTTGTCGACCATCGCATTCGATTCAGATT 591  
 QY 702 GTGAGCCACGACAAAGATTACAAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAG 758  
 DB 592 TTGAGCCATGACAAAGATTACAAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAG 651  
 QY 759 CATGAGTGTCAAGGAGGCGCAAGTAA 785  
 DB 652 TCTATGCTGCGGAGTAAGGCGCAAGTAA 678

RESULT 15  
 US-10-492-081B-2  
 ; Sequence 2, Application US/10492081B  
 ; Publication NO. US20050106661A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MIYAWAKI, Atsushi  
 ; APPLICANT: KARASAWA, Satoshi  
 ; APPLICANT: ARAKI, Toshio  
 ; TITLE OF INVENTION: Fluorescent Protein  
 ; FILE REFERENCE: P25141  
 ; CURRENT APPLICATION NUMBER: US/10/492,081B  
 ; CURRENT FILING DATE: 2004-04-09  
 ; PRIOR APPLICATION NUMBER: PCT/JP02/10529  
 ; PRIOR FILING DATE: 2001-10-11  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 2  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: Galaxea fascicularis  
 US-10-492-081B-2

Query Match 45.6%; Score 486.4; DB 9; Length 678;  
 Best Local Similarity 84.8%; Pred. No. 1e-135;  
 Matches 560; Conservative 0; Mismatches 91; Indels 9; Gaps 1;  
 QY 102 AAGGGTGTGATTAAACACAGATGAAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 161



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 21:14:53 ; Search time 666 Seconds  
(without alignments)  
6403.292 Million cell updates/sec

Title: US-10-757-356-17

Perfect score: 1066

Sequence: 1 attcgccctggtgattgga.....atcggaagcgaataaaaaa 1066

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:  
1: /SID55/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
2: /SID55/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
3: /SID55/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
4: /SID55/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
5: /SID55/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
6: /SID55/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
7: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
8: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
9: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
10: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
11: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
12: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
13: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
14: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
15: /SID55/ptodata/1/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.6	24.1	678	8	US-10-209-208-2
2	256.6	24.1	678	14	US-11-218-880-2
3	256.6	24.1	678	14	US-11-100-988-1
4	252	23.6	678	14	US-11-082-154A-39
5	230	21.6	678	14	US-11-100-988-4
6	226	21.2	645	11	US-11-208-927-1
7	223.6	21.0	678	8	US-10-209-208-5
8	223.6	21.0	678	14	US-11-218-880-5
9	219.8	20.6	678	11	US-11-187-622-7
10	219.8	20.6	681	8	US-10-209-208-3
11	219.8	20.6	681	14	US-11-218-880-3
12	219.8	20.6	681	14	US-11-218-880-23
13	219.8	20.6	4692	14	US-11-082-154A-29
14	219.8	20.2	678	11	US-11-187-622-13
15	215.8	20.2	675	11	US-11-187-622-14
16	215.8	20.2	675	11	US-11-205-740-3
17	215.8	20.2	1635	11	US-11-205-740-3
18	215.8	20.2	1635	11	US-11-205-740-9

19	215.8	20.2	1635	11	US-11-205-740-11
20	215.8	20.2	6706	8	US-10-655-872-4
21	215.8	20.2	7927	8	US-10-655-872-7
22	214.2	20.1	1815	11	US-11-205-740-4
23	214.2	20.1	1815	11	US-11-205-740-10
24	214.2	20.1	1815	11	US-11-205-740-12
25	214.2	20.1	2361	11	US-11-205-740-1
26	214.2	20.1	2361	11	US-11-205-740-7
27	214.2	20.1	2625	11	US-11-205-740-2
28	214.2	20.1	2625	11	US-11-205-740-8
29	212.4	19.9	681	8	US-10-209-208-7
30	212.4	19.9	681	14	US-11-218-880-7
31	207	19.4	678	11	US-11-187-622-11
32	203.2	19.1	687	11	US-11-187-622-1
33	199.6	18.7	678	8	US-10-209-208-9
34	199.6	18.7	678	14	US-11-218-880-9
35	196.4	18.4	678	8	US-10-209-208-80
36	188.4	17.7	690	11	US-11-187-622-19
37	177.8	16.7	865	11	US-11-187-622-5
38	176.2	16.5	693	11	US-11-187-622-3
39	171	16.0	705	11	US-11-187-622-15
40	165	15.5	696	11	US-11-187-622-9
41	164.8	15.5	705	11	US-11-187-622-17
42	159	14.9	654	11	US-11-187-622-23
43	157.2	14.7	707	11	US-11-187-622-21
44	132.2	12.4	720	11	US-11-149-015-1
45	126	11.8	720	11	US-11-149-015-3

## ALIGNMENTS

RESULT 1  
US-10-209-208-2  
; Sequence 2, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: misc:feature  
; LOCATION: (1)...(678)  
; OTHER INFORMATION: wild-type DeRed  
US-10-209-208-2

Query Match	24.1%	Score	256.6	DB	8	Length	678
Best Local Similarity	62.8%	Pred. No.	3.4e-61				
Matches	419	Conservative	0	Mismatches	239	Indels	9
Gaps	1						
Qy	102	AAGGGTGTGATTAAACCAAGCATCAAGTGAAGTGGTATGGAAGGTGCTGTAACGGG	161				
Db	13	AAGAAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTCCATGGAAGAACCGTCAATGGG	72				
Oy	162	CACAAGTTCGTGGTTGAAGGAGATGGAAGGAGCCCTTTCCAGCGGAACACAGACTATG	221				



```

; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: wild-type red fluorescent protein based on Discosoma species
US-11-100-988-1

```

```

Query Match      24.1%; Score 256.6; DB 14; Length 678;
Best Local Similarity 62.8%; Pred. No. 3.4e-61;
Matches 419; Conservative 0; Mismatches 239; Indels 9; Gaps 1;

QY 102 AAGGCTGTGATTAAACACAGATCAAGATGAAGCTGGTATGAAGTGTCTGTAACAGGG 161
Db 13 AAGAAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTCGCATGGAAGGACGGTCAATGGG 72

QY 162 CACAAGTTCCTGGTTGAAGGAGATGGAAGGAGGCTTTTCAGCGAACACAGACTATG 221
Db 73 CACGAGTTTGAATAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 132

QY 222 GACCTTTACAGTATAGAGGCGCACCATTCCTTTTGCCTTACGATATCTTGACAAAGTA 281
Db 133 AAGCTTAAAGTAAACCAAGGCGGACCTTTGSCATTTGCTTGGATATTTGTCAACACAA 192

QY 282 TTGATTAAGCAACAGGTTTTCGCAAAATACCCAGAGATACAGATAGATATTTCAAG 341
Db 193 TTTTCAATGGAAGCAAGGATATATGTCAGCACCTTCCGACATACAGACTATAAAG 252

QY 342 CAGACCTTTCCTGAGGGTACTTCTGGACGAGGATGATGATGATGATGATGATGATGAT 401
Db 253 CTGTCATTTCTGAGGAGTTTAAATGGGAAAGGATGATGATGATGATGATGATGATGATG 312

QY 402 TGATCGCCCAACACAGATACAAATGATGGAAGGCGTGCAGACTGTTTTCCTTATAA 461
Db 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGATGG-----CTGTTTCATCTACAG 363

QY 462 ATTGAGTTTGAATGAGTGTGAACTTCTGCGCAATGCTGCTGCTGCTGCTGCTGCTGCT 521
Db 364 GTCAGTTTCAATGTCGCTGAACTTCTTCCGATGAGACCTGTTATGCAAAAGAGACAATG 423

QY 522 AAATGGAGCCATCCACTGAGATTAATGATGCGGCTGATGAGTGTGAGGCTGATGTT 581
Db 424 GGTGGGAGCCAGCAGTACGCTTGTATCTCTGCTGATGGGCTGTTGAAAGGAGAGATTT 483

QY 582 AACATGGCTCTGTGTCTGAAGGAGGTGGCCATTACCGATGTGACTTCAAAACTACTTAC 641
Db 484 CATAAGGCTCTGAAGCTGAAGACGCTGGTCACTTACTAGTTGAATTCAAAGTATTAC 543

QY 642 AAAGTAAAGAGTTTCTGCTGGTGGCCAGACTATCACTTTGTGACCATCGCATTTGAGATT 701
Db 544 ATGGCAAGAGGCTGTGAGCTTACAGGCTACTACTATGTTGACTCCAAACTGGATATA 603

QY 702 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGACGAGCATGCCGAAGCTCGTCAT 761
Db 604 ACAAGCCACAAGAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGAGCCAC 663

QY 762 GGACTGT 768
Db 664 CATCTGT 670

```

```

RESULT 4
US-11-082-154A-39
; Sequence 39, Application US/11082154A
; Publication No. US20060024820A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS

```

```

; FILE REFERENCE: 17084-022003 (420C)
; CURRENT APPLICATION NUMBER: US/11/082,154A
; CURRENT FILING DATE: 2005-03-15
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 10/161,403
; PRIOR FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Discosoma species
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45)...(737)
; OTHER INFORMATION: Nucleotide sequence encoding red fluorescent
; OTHER INFORMATION: protein (FP593)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AF272711
; DATABASE ENTRY DATE: 2000-09-26
US-11-082-154A-39

```

```

Query Match      23.6%; Score 252; DB 14; Length 876;
Best Local Similarity 62.9%; Pred. No. 7.3e-60;
Matches 428; Conservative 0; Mismatches 240; Indels 12; Gaps 2;

QY 102 AAGGCTGTGATTAAACACAGATCAAGATGAAGCTGGTATGAAGTGTCTGTAACAGGG 161
Db 57 AAGATGTGATCAAGGAGTTTCATGAGTTCAAGTTTCGATGGAAGGACGGTCAATGGG 116

QY 162 CACAAGTTTCTGCTGGTTGAAGGAGATGGAAGGAGGAGCTTTTCGACGGAACACAGATG 221
Db 117 CACGAGTTTGAATATAAAGGCGAAGGTGAAGGAGGAGCTTACGAGGTCACCTTCCGTA 176

QY 222 GACCTTACAGTATAGAGGCGCACCATTCCTTTCGCTTACGATCTTGACAAAGTA 281
Db 177 AAGCTTATGTTAAACCAAGGCTGGACCTTTGCCATTTGCTTTTGTATATTTTGTCAACAA 236

QY 282 TTCGATTACGCGCAACAGGCTATTCGCCAAATACCCAGAGAGACATAGCAGATTATTTCAAG 341
Db 237 TTTTCAAGTATGGAAGCAAGGATATATGTCACAAACCTCTGCCGACATACAGACTATAAAG 296

QY 342 CAGAGCTTCTGAGGGGTACTTCTTGGGAACCAAGATGACATACGAAAGCCAGGCAAT 401
Db 297 CTGTCATTTCTGAGGGATTTAAATGGGAAAGGCTCATGAACTTTGAAGACGCTGCGGTG 356

QY 402 TGCATCGCCACAAACAGACATACAAATGATGGAAGGCTGCAGACTGTTTTCCTTATAA 461
Db 357 GTTACTGTATCCCAAGATTCAGATTGA-----AAGACGGCTGTTTTCATCTACGAG 407

QY 462 ATTCGATTGTGCTGTGAACTTTCCTCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 521
Db 408 GTCAAGTTTCATTTGGGCTGAACTTCTTCTGATGACCTGTTTATGAGAGGAGGACACGG 467

QY 522 AAATGGAGCCATCCACTGAGATTAATGATGCGCTGATGGAAGTGTGAGGCTGATGTT 581
Db 468 GGCTGGGAAGCCAGCTCTGAGCGTTTGTATCTCTGATGGGTGCTGAAAGGAGACATC 527

QY 582 AACATGGCTCTGTGCTTGAAGGAGGTCGCCATTAACGATGTCATCTCAAACTACTTAC 641
Db 528 CATATGGCTCTGAGGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 587

QY 642 AAAGTAAAGAGGTTT---GTCCGGTTGCGAGACTATCACTTTGTGGACCATCGCATTTGAG 698
Db 588 ATGGTAAAGAGGCTTCAGTGCAGTTGCGAGCTACTATTATGTTGACTCCAACTGGAT 647

QY 699 ATTGTGAGCCACGACAAAGATTACAAAGGTTAAGCTGACAGCATGCCGAAGCTCGT 758
Db 648 ATGACGAGCCACAAAGAGATTACACAGTCGTTGAGCAGTATGAAAGGAGGAGGAGGAGG 707

```

```
QY 759 CATGACTGTCAAGGAGGC 778
|||
Db 708 CACCATCCGTTCAATTAAGCC 727

RESULT 5
US-11-100-988-4
; Sequence 4, Application US/11100988
; Publication No. US2006000878A1
; GENERAL INFORMATION:
; APPLICANT: Gluck, Benjamin S.
; APPLICANT: Bevis, Brooke
; APPLICANT: Strongin, Daniel E.
; APPLICANT: Baker, David
; APPLICANT: Scalley-Kim, Michelle
; TITLE OF INVENTION: MONOMERIC RED FLUORESCENT PROTEINS
; FILE REFERENCE: 092234-9002-US01
; CURRENT APPLICATION NUMBER: US/11/100,988
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: US 60/560,340
; PRIOR FILING DATE: 2004-04-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mutant red fluorescent protein DrRed.M1 based on Discosoma
; OTHER INFORMATION: species
; NAME/KEY: CDS
; LOCATION: (1)..(678)
US-11-100-988-4

Query Match 21.6%; Score 230; DB 14; Length 678;
Best Local Similarity 60.6%; Pred. No. 9.1e-54;
Matches 399; Conservative 0; Mismatches 250; Indels 9; Gaps 1;

QY 103 AGGCTGTATTAACGACGATGAAGTGAAGTGTGTAAGGCTGTAAACGGCC 162
Db 14 AGGACGTATCAAGGAGTTCAAGTGTGCGATGAGGCTCCGTTGAAGCC 73

QY 163 ACAAGTTCGTGTGAAGGAGATGAAAGGAAAGCTTTTCACGGAACACAGACTATGG 222
Db 74 ACTACTTCAGATTCAGGGCGAGGGGAGGCGGCTACGAGGCGCCACGACCGCCA 133

QY 223 ACCTTACAGTCTAGAAGCGCACCATTTGCCCTTTTCGCTTACGATATCTTGAACAAGTAT 282
Db 134 AGCTGAGGTGACCAAGGCGGCCCTTGCCTTGGGACATCTCTGCCCCCACT 193

QY 283 TCGATTAGGCAACAGGATTTTCGCAATATCCAGAGACATAGACAGATTTTCAAGC 342
Db 194 TCCAGTACGGCTTCAAGGCTACGTGAAGCACCCTCCGCGACATCCCGCACTACATGAAGC 253

QY 343 AGACGTTTCTCAGGGTACTTCTGGGAACGAAGATGACATACGAACACGAGGCACTT 402
Db 254 TGTCTTCTCCGAGGGCTTCACTGGAGCGCTCCATGAATTTGAGACGCGCGGTGG 313

QY 403 GCATGCCACAAACGACATTAACAATGATGGAAGGGCTGCAAGCACTGTTTTCCTATAAAA 462
Db 314 TGGAGGTGAGCAGGACTCTCTCCCTGCAGAGCGGCACC-----TTCACTACAAGG 364

QY 463 TTCGATTGATGGTGTGAATCTTCTGCAATATGGTTCAGTTATGAGAGGAAGACGCTGA 522
Db 365 TGAAGTTCAAGGGCGTGAATCTTCCCGCGCGAGCGGCCCTTAATGAGAGGAAGACTGCCG 424

QY 523 AATGGAGCCATCCACTGAGTAATCTATGCGCGTGTAGTGTGCTGAGGGGTGATGTTA 582
Db 425 GCTGGAGGCCCTTCCACCGGAGAGCTGTACCCCGAGGACGGCGTGTGAGGGCGGAGATCT 484

QY 583 ACATGGCTCTGTGTTGAAGGAGGTGGCCATTACCGATGTGACTTGTCAAACTACTTACA 642
|||
```

```
Db 485 CCCACGCGCTGAAGCTGAAGGACGGCGCCACTACACCTGCGACTTCAAGACCGTGTACA 544
QY 643 AAGCTAAGAAGTTGTCGGTTGCCAGACTATCACTTTGTGGACCATCGCATTTGAGATTG 702
Db 545 AGGCCAAGAAGCCGTCGAGCTGCCCGCAACCACTACGTGGACTCCAAAGCTGGACATCA 604
QY 703 TGAGCCACGACAAAGATTACAAAGGTTAAGCTGCACGAGCATGCCGAAGCTCCGTCA 760
Db 605 CCAACCAACAGAGGACTACACCGTGTGGAGCAGTACGAGCAGCGCCGAGGCCGCCA 662

RESULT 6
US-11-208-927-1
; Sequence 1, Application US/11208927
; Publication No. US20060063229A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yi
; APPLICANT: Falkowski, Paul
; TITLE OF INVENTION: mmp encoding nucleic acids,
; FILE REFERENCE: polypeptides, antibodies and methods of use thereof
; FILE REFERENCE: Ruc.02-099US
; CURRENT APPLICATION NUMBER: US/11/208,927
; CURRENT FILING DATE: 2005-08-22
; PRIOR APPLICATION NUMBER: US/10/652,529
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: 60/407,478
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Meandrina meandrites
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (329)...(329)
; OTHER INFORMATION: n= g or a
US-11-208-927-1

Query Match 21.2%; Score 226; DB 11; Length 645;
Best Local Similarity 62.8%; Pred. No. 1.2e-52;
Matches 409; Conservative 0; Mismatches 221; Indels 21; Gaps 3;

QY 112 TTAACACAGACATGAAGTGAAGCTGCGTATGGAAGTGTGTAAACGGGCAACAGTTTCG 171
Db 8 TTCAAACGCAAGTGAAGTGAATATCTATGATGGCACTTCAACGGGCAAGCTTTA 67

QY 172 TGGTTGAAGGAGATGGAAGGAGCTTTTCGACGGAAACACAGACTATGACCTTACAG 231
Db 68 CTGTCGTAGTGAAGGACCGGCAATCCATACGAGGGACATCAGTCGTTAAACCTCACCG 127

QY 232 TCATAGAAGGCGCACCATTTGCTTTTCGTTACGATATCTTGCACACAGTATTCGATTACG 291
Db 128 TCA---AGGCGGAGCCCTCTGCTTTGCGGTTTGACATATTTGTCAGAACGTTTCACTTATG 184

QY 292 GCAACAGGGTATTCGCGCAATATACCAAGAAAGCATAGCAGATTTTTCAGCAGACGTTTC 351
Db 185 GCAACAGGGTATTTTACTAAGTACCTTGAAGGAAACCGCACTATTTCAAGGAGCATTTTC 244

QY 352 CTGAGGGGTACTTCTGGGAACGAGCATGACATAGGAAGCAAGGGCATTTTGCATCGCCA 411
Db 245 CTGAGGGCTGACATGGAAGGCAATGACGTTTGAAGATGGCGGCAATTTGCACAGTCG 304

QY 412 CAACAGCATATCAATGATGAAGGCGTCGACGACTGTTTTCCTATATAAATTCGATTG 471
Db 305 CAGCGAAATTAAGCCTTACAGGAAGGCTCTGAGCAAGGCTCTGAGCAAGGCTTCAAAATTTG 355

QY 472 ATGCTGTGAATCTTCTGCGCAATGCTCCAGTTATGAGAGGAAGACGCTGAAATCGGAGC 531
Db 356 TCGCGGTGAATCTTCCGCGCAATGAGCCTGTAAATCCAGAGAGACGCTTGGCTGGGAGA 415

QY 532 CATCCACTGAGATAATGTATGCGCGTGTATGAGGCTGATGAGGCTGATTTAACTGGCTC 591
|||
```

Db 416 CATCTACCGAGAAATGGCTGCCAACGATGATCCGTACAGGCGTATGATACCATGTTCC 475  
QY 592 TGTGTTGCTTGAAGAGGTGGCCATTTACCGATGTGACTTCAAACTACTTACAAAGCTAAGA 651  
Db 476 TCAAGCTGGAGGAGGTGGCGGTCACAAAATGCTATTTCGTAACTAATCAAAAGCCAAA 535  
QY 652 AGTTTGCCTGGTTCGCCAGACTATCACTTTGTGGACCATCGCATTTGAGATTGTGAGCCACG 711  
Db 536 GGGCCGTGAAGGTGCCAGATAACCATTTTGTGTGGCATCGCTCTCAGGAATGGC---- 591  
QY 712 ACAAGATTACACAAGATTAACTGATGCGACGAGCATGCCGAAGCTCGTCATG 762  
Db 592 -----GATGCAACACACTGTTGAGCTGGAAGAACTGCAGAAGCTCGTTAG 637

## RESULT 7

US-10-209-208-5  
; Sequence 5, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "T1"  
US-10-209-208-5

Query Match 21.0%; Score 223.6; DB 8; Length 678;  
Best Local Similarity 59.8%; Pred. No. 5.6e-52;  
Matches 398; Conservative 0; Mismatches 259; Indels 9; Gaps 1;  
QY 103 AGGGTGTGATTAAACACAGACATGAGATGAAGTGGTATGGAAGTGTCTGTAACGGGC 162  
Db 14 AGGAGTCTACAGGAGTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTGAACGGCC 73  
QY 163 ACAAGTTCGTGTGAAGGAGATGGAAGGAAAGGAGCCCTTTGACGCGAACAAGATATGG 222  
Db 74 ACGAGTTCGAGATCGAGGCGAGGGCGGCGGCCCTTACGAGGCCACCGACCGCA 133  
QY 223 ACCTTACAGTCTAGAGGCGCACCATTCCTTTGCTTACGATATCTTGAACAAGTAT 282  
Db 134 AGCTGAAGTGAACCAAGGCGCGCCCTTGCCTTGGGACATCTCTGCCCCCAGT 193  
QY 283 TCGATTACGGCACAGGGTATTTCGCCAATATCCAGAGACATAGCAGTATTTCAAGC 342  
Db 194 TCCAGTACGGCTCCCAAGGTGTACGTGAAGCAACCCGCGACATCCCGACTACAAGAGC 253  
QY 343 AGACGTTTCTGAGGGTACTTCTGGGAACGAAGATGATACATACGAAGACACAGGCGATT 402  
Db 254 TGTCTTCCCCAGGGCTTCAAGTGGAGCGGTGATGAACTTCGAGAGCGCGCGTGG 313  
QY 403 GCATCGCCACAAACGACATACAAATGATGAAGGCGGTGCAAGCTGTTTTCCTTATAAA 462  
Db 314 TGACCGTGACCCAGGACTCTCTCCCTGCAG-----GACGGCTCTTCTACTCAAGG 364

QY 463 TTCGATTTCATGCTGTGAACCTTTCTCTGCCAATGCTCCAGTTATGCGAGAGCAACGCTGA 522  
Db 365 TGAAGTTCATCGCGTGAACCTTTCCCTCCGACGCCCCCGTAAATGCAAGAAAGACTATGG 424  
QY 523 AATGGAGCCATCCACTGAGATAAATGATGCGCGTGTATGGAGTGTCTGAAGGGTGTATGTTA 582  
Db 425 GCTGGAGGCGCTCCACCGAGCGCTGTATCCCGCGACGCGTGTGTAAGGGCGAGATCC 484  
QY 583 ACATGGCTCTGTTGTTGAAGGAGGTGGCCATTAACGATGTGACTTCAAAACTACTTACA 642  
Db 485 ACAAGGCCCTGAAGCTGAAGGACGCGCGCCACTTACCTGGTGGAGTTCAAGTCCATCTACA 544  
QY 643 AAGCTAAGAAGGTGTGTCGGTTGCCAGACTATCACTTTGTGGACCATCGCATTTGAGATTG 702  
Db 545 TGCCCAAGAGCGCGTGGAGCTGCCGGCTACTACTACGTGGACTCCAAGCTCGACATCA 604  
QY 703 TGAGCCACGACAAAGATTACAAAGTTAAGTGCACGAGCATGCCGAAGCTTCGTCTATG 762  
Db 605 CCTCCACAAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCGCGAGGGCGCCACC 664  
QY 763 GACTGT 768  
Db 665 ACCTGT 670

## RESULT 8

US-11-218-880-5  
; Sequence 5, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CPI  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "T1"  
US-11-218-880-5

Query Match 21.0%; Score 223.6; DB 14; Length 678;  
Best Local Similarity 59.8%; Pred. No. 5.6e-52;  
Matches 398; Conservative 0; Mismatches 259; Indels 9; Gaps 1;  
QY 103 AGGGTGTGATTAAACACAGACATGAGATGAAGTGGTATGGAAGTGTCTGTAACGGGC 162  
Db 14 AGGAGTCTACAGGAGTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTGAACGGCC 73  
QY 163 ACAAGTTCGTGTGAAGGAGATGGAAGGAAAGGAGCCCTTTGACGCGAACAAGATATGG 222  
Db 74 ACGAGTTCGAGATCGAGGCGAGGGCGCGGCCCTTACGAGGCCACCGACCGCA 133  
QY 223 ACCTTACAGTCTAGAGGCGCACCATTCCTTTGCTTACGATATCTTGAACAAGTAT 282  
Db 134 AGCTGAAGTGAACCAAGGCGCGCCCTTGCCTTGGGACATCTCTGCCCCCAGT 193  
QY 283 TCGATTACGGCAACAGGGTATTTCGCCAATATCCAGAGACATAGCAGTATTTCAAGC 342  
Db 194 TCCAGTACGGCTCCCAAGGTGTACGTGAAGCAACCCGCGACATCCCGACTACAAGAGC 253





```

Query Match      20.6%; Score 219.8; DB 8; Length 681;
Best Local Similarity 59.4%; Pred. No. 6.4e-51;
Matches 396; Conservative 0; Mismatches 262; Indels 9; Gaps 1;

QY 102 AAGGTTGTGATTAACACAGACATGAGATGAAGTGGTATGAAGTGTCTGTAACGGG 161
Db 16 AAGAAAGTGTATCAAGGAGTTTCATGCGCTTCAAGGTGCGATGGAGGCGCCGTAACGGC 75
QY 162 CACAAGTTCTGTGTTGAAGGAGATGGAAGGAGATGGAAGGAGATGGAAGGAGATG 221
Db 76 CACGAGTTTCAGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 135
QY 222 GACCTTACAGTCAATAGAGGCGCACCATTCCTTTTCGCTTACGATATCTTGACAAAGTA 281
Db 136 AAGCTGAAAGTGCACCAAGGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 195
QY 282 TTCGATTACGGCAACAGGGTATTCGCCAAATACCCAGAGACATAGCAGATTATTTCAAG 341
Db 196 TTCAGATGCGGCTCCCAAGGTGTACGTGAAGCACCCCGCGGACATCCCGGACTACAGAAG 255
QY 342 CAGACGTTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATAGCAAGCAGGCGCAT 401
Db 256 CTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAGACCGCGCGTG 315
QY 402 TGCATCGCCACAAACGACATAACAATGATGGAAGGCGTGCAGCAGTGTTCCTTATAAA 461
Db 316 GTGACCGTGACCCAGGACTCTCTCCCTGCAG-----GACGGCTGCTTCATCTACAAG 366
QY 462 ATTGATTGATGGTGTGAACCTTCTGCGCAATGTCGCAATGTCGAGAGGAGAGCGTG 521
Db 367 GTGAAGTTTCATCGGCGTGAACCTTCCCTCCGACGCGCGCTAATGCAAGAAGACCATG 426
QY 522 AATGGAGGCGCATCCACTGAGATAATGATGCGGTGATGGAGTGTCTGAAGGTTGATGT 581
Db 427 GGTGGAGGCGCTCCACGAGCGCTGTACCCCGCGAGCGGTGCTGAAGGCGGAGATC 486
QY 582 AACATGGCTCTGTGCTGAAGGAGTGGCCATTTACCGATGTGACTTTCAAAACTACTTAC 641
Db 487 CACAGGCGCTGAAGTGAAGGAGCGGCGCCACTACTCTGGTGGAGTTCAAGTCCATCTAC 546
QY 642 AAGCTAAGAGGTTCTCGGTTCGAGACTATCACTTTGTGACCATCGCATTTGAGATT 701
Db 547 ATGGCCAAAGACCCGTCAGCTGCGCGCTACTACTAGTGGACTCCCAAGCTGGACATC 606
QY 702 GTGAGCCAGCAAAAGATTACAAAGGTTAAGCTGACGAGCATGCCGAGCTCGTCAT 761
Db 607 ACCTCCCAACAGGAGTACACCATCTGTGGAGCAGTACGAGCGCACCGAGGCGGCCAC 666
QY 762 GGACTGT 768
Db 667 CACCTGT 673

```

```

RESULT 11
US-10-209-208-23
; Sequence 23, Application US/10209208
; Publication No. US2005024921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80

```

```

; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DeRed with mammalian
; OTHER INFORMATION: codon usage
US-10-209-208-23

Query Match      20.6%; Score 219.8; DB 8; Length 681;
Best Local Similarity 59.4%; Pred. No. 6.4e-51;
Matches 396; Conservative 0; Mismatches 262; Indels 9; Gaps 1;

QY 102 AAGGTTGTGATTAACACAGACATGAGATGAAGTGGTATGAAGTGTCTGTAACGGG 161
Db 16 AAGAAAGTGTATCAAGGAGTTTCATGCGCTTCAAGGTGCGATGGAGGCGCCGTAACGGC 75
QY 162 CACAAGTTCTGTGTTGAAGGAGATGGAAGGAGATGGAAGGAGATGGAAGGAGATG 221
Db 76 CACGAGTTTCAGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 135
QY 222 GACCTTACAGTCAATAGAGGCGCACCATTCCTTTTCGCTTACGATATCTTGACAAAGTA 281
Db 136 AAGCTGAAAGTGCACCAAGGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 195
QY 282 TTCGATTACGGCAACAGGGTATTCGCCAAATACCCAGAGACATAGCAGATTATTTCAAG 341
Db 196 TTCAGATGCGGCTCCCAAGGTGTACGTGAAGCACCCCGCGGACATCCCGGACTACAGAAG 255
QY 342 CAGACGTTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATAGCAAGCAGGCGCAT 401
Db 256 CTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAGACCGCGCGTG 315
QY 402 TGCATCGCCACAAACGACATAACAATGATGGAAGGCGTGCAGCAGTGTTCCTTATAAA 461
Db 316 GTGACCGTGACCCAGGACTCTCTCCCTGCAG-----GACGGCTGCTTCATCTACAAG 366
QY 462 ATTGATTGATGGTGTGAACCTTCTGCGCAATGTCGCAATGTCGAGAGGAGAGCGTG 521
Db 367 GTGAAGTTTCATCGGCGTGAACCTTCCCTCCGACGCGCGCTAATGCAAGAAGACCATG 426
QY 522 AATGGAGGCGCATCCACTGAGATAATGATGCGGTGATGGAGTGTCTGAAGGTTGATGT 581
Db 427 GGTGGAGGCGCTCCACGAGCGCTGTACCCCGCGAGCGGTGCTGAAGGCGGAGATC 486
QY 582 AACATGGCTCTGTGCTGAAGGAGTGGCCATTTACCGATGTGACTTTCAAAACTACTTAC 641
Db 487 CACAGGCGCTGAAGTGAAGGAGCGGCGCCACTACTCTGGTGGAGTTCAAGTCCATCTAC 546
QY 642 AAGCTAAGAGGTTCTCGGTTCGAGACTATCACTTTGTGACCATCGCATTTGAGATT 701
Db 547 ATGGCCAAAGACCCGTCAGCTGCGCGCTACTACTAGTGGACTCCCAAGCTGGACATC 606
QY 702 GTGAGCCAGCAAAAGATTACAAAGGTTAAGCTGACGAGCATGCCGAGCTCGTCAT 761
Db 607 ACCTCCCAACAGGAGTACACCATCTGTGGAGCAGTACGAGCGCACCGAGGCGGCCAC 666
QY 762 GGACTGT 768
Db 667 CACCTGT 673

```

```

RESULT 12
US-11-218-880-3
; Sequence 3, Application US/11218880
; Publication No. US20060003420A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI

```

```
; CURRENT APPLICATION NUMBER: US/11/218,880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
; OTHER INFORMATION: codon usage
US-11-218-880-3

Query Match          20.6%; Score 219.8; DB 14; Length 681;
Best Local Similarity 59.4%; Pred. No. 6.4e-51;
Matches 396; Conservative 0; Mismatches 262; Indels 9; Gaps 1;

QY 102 AAGGCTGTGATTAAACACGACATGAAGTGAAGTGCCTATGGAAGTGTCTGTAACCGG 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 AAGAACGTCATCAAGGAGTTCATCGCTTCAAGGTGCGCATGGAGGCGCACCGTGAACGC 75

QY 162 CACAAGTTTCGTGTTGAAGGAGATGGAAGAGGGAAGCCTTTTCGACGGAACAACAGACTATG 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 CACGAGTTCGAGATCGAGGGCGAGGGCGGCGCCCTACGAGGGCCACACACCGTG 135

QY 222 GACCTTACAGTCATAGAGGGCCACCATTTGCCCTTTCGCTTACGATATCTTGACACAGTA 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 AAGCTGAAGTGAACCAAGGGCGGCCCTTGCCTTTCGCTTGGGACATCTGTCCCCCAG 195

QY 282 TTCGATTACGGCAACAGGGTATTCGCCAATAACCCAGAACATAGCAGATTTATTTCAAG 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 TTCAGTACGGCTCCAAAGGTGTCGTAAGACACCCCGCCGACATCCCGACATCAAGAG 255

QY 342 CAGACGTTTCGTGAGGGTACTTCTGGGNAACGAGCATGACATAGACAGACAGGCGATT 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 CTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGAATTCGAGGACGGCGCGTG 315

QY 402 TGCATCGCCACAAACGACATACATATGATGGAAGCGTCGACGACTGTTTTCCTATATAA 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 GTGACCGTGACCCAGGATCTCCTCCCTGCAG-----GACGGTGTCTTATCTACAG 366

QY 462 ATTCGATTGATGTGTGAACCTTTCCTGCCAATGGTCCAGTTATGACAGGAAGACGCTG 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGCGCCGTAATGACAGGAAGACCATG 426

QY 522 AAATGGGAGCCATCCACTGAGATAATGTATGCGCGTGAATGAGGTGCTGAAGGGTATGTT 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 GGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGACGCGGTGCTGAAGGGCGAGATC 486

QY 582 AACATGGCTCTGTGTTGAAGGAGTGGCCATTACCGATGTGACTTCAAACTACTTAC 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 CACAAGGCCCTGAAGCTGAAGGACGCGCGCCACTACCTGTGGAGTTCAAGTCCATCTAC 546

QY 642 AAAGCTAAGAAGGTTGTCGGTTGCCAGACTATCACTATTTGTGGACCATCGCATTTGAGATT 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 ATGGCCAAAGAGCCCGTGCAGCTGCCCGGTACTACTACGTGGACTCCAAAGCTGGACATC 606

QY 702 GTGACCCAGCAAAAGATTACAAAGGTTAAGCTGACGAGCATGCGGAAGCTCGTAT 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 ACCTCCCAACAGGAGTACTACACCATCTGTGGAGCAGTACGAGCGCACCGGCGCCAC 666

QY 762 GGACTGT 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 CACCTGT 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 13

```
US-11-218-880-23
; Sequence 23, Application US/11/218880
; Publication No. US20060003420A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/11/218,880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
; OTHER INFORMATION: codon usage
US-11-218-880-23
```

```
Query Match          20.6%; Score 219.8; DB 14; Length 681;
Best Local Similarity 59.4%; Pred. No. 6.4e-51;
Matches 396; Conservative 0; Mismatches 262; Indels 9; Gaps 1;

QY 102 AAGGCTGTGATTAAACACGACATGAAGTGAAGTGCCTATGGAAGTGTCTGTAACCGG 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 AAGAACGTCATCAAGGAGTTCATCGCTTCAAGGTGCGCATGGAGGCGCACCGTGAACGC 75

QY 162 CACAAGTTTCGTGTTGAAGGAGATGGAAGAGGGAAGCCTTTTCGACGGAACAACAGACTATG 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 CACGAGTTCGAGATCGAGGGCGAGGGCGGCGCCCTACGAGGGCCACACACCGTG 135

QY 222 GACCTTACAGTCATAGAGGGCCACCATTTGCCCTTTCGCTTACGATATCTTGACACAGTA 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 AAGCTGAAGTGAACCAAGGGCGGCCCTTGCCTTTCGCTTGGGACATCTGTCCCCCAG 195

QY 282 TTCGATTACGGCAACAGGGTATTCGCCAATAACCCAGAACATAGCAGATTTATTTCAAG 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 TTCAGTACGGCTCCAAAGGTGTCGTAAGACACCCCGCCGACATCCCGACATCAAGAG 255

QY 342 CAGACGTTTCGTGAGGGTACTTCTGGGNAACGAGCATGACATAGACAGACAGGCGATT 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 CTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGAATTCGAGGACGGCGCGTG 315

QY 402 TGCATCGCCACAAACGACATACATATGATGGAAGCGTCGACGACTGTTTTCCTATATAA 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 GTGACCGTGACCCAGGATCTCCTCCCTGCAG-----GACGGTGTCTTATCTACAG 366

QY 462 ATTCGATTGATGTGTGAACCTTTCCTGCCAATGGTCCAGTTATGACAGGAAGACGCTG 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGCGCCGTAATGACAGGAAGACCATG 426

QY 522 AAATGGGAGCCATCCACTGAGATAATGTATGCGCGTGAATGAGGTGCTGAAGGGTATGTT 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 GGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGACGCGGTGCTGAAGGGCGAGATC 486

QY 582 AACATGGCTCTGTGTTGAAGGAGTGGCCATTACCGATGTGACTTCAAACTACTTAC 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 CACAAGGCCCTGAAGCTGAAGGACGCGCGCCACTACCTGTGGAGTTCAAGTCCATCTAC 546

QY 642 AAAGCTAAGAAGGTTGTCGGTTGCCAGACTATCACTATTTGTGGACCATCGCATTTGAGATT 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 ATGGCCAAAGAGCCCGTGCAGCTGCCCGGTACTACTACGTGGACTCCAAAGCTGGACATC 606

QY 702 GTGACCCAGCAAAAGATTACAAAGGTTAAGCTGACGAGCATGCGGAAGCTCGTAT 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 ACCTCCCAACAGGAGTACTACACCATCTGTGGAGCAGTACGAGCGCACCGGCGCCAC 666

QY 762 GGACTGT 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 CACCTGT 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



Db	273	CAAGTGGGAGCGGTGATGAACCTTCGAGACGGGGCGTGGCGACCGTGACCCGAGGACTC	332
Qy	422	AACAAATGATGAAGGCGTCGACGACTGTTTTCCTATAAAATTCGATTTGATGGTGTGAA	481
Db	333	CTCCCTGCAG-----GACGGCTGCTTCATCTACAAGGTGAAGTTTCATCGGCGTGAA	383
Qy	482	CTTTCCTGCCAATGGTCCAGTTATGCAGAGGAGAACGCTGAATGGGAGCCATCCACTGA	541
Db	384	CTTCCCTCCGACGCGCCCGGTGATGCAGAGAGACCATGGGCTGGGAGGCGCTCCACCGA	443
Qy	542	GATTAATGTATGGCGGTGATGGAGTGTGAAGGGTGTGTTAAATGCTCTGTGCTTGA	601
Db	444	CGGCTGTACCCCGCGGACGGCGTGTGAAGGGGAGACCCACAGGCCCTGAAGCTGAA	503
Qy	602	AGGAGTGGCCATTACCGATGTGACTTCAAACTACTTACAAAGCTAAGAGGTTGTCCG	661
Db	504	GGACGGGGCCACTACTGTGTGGAGTTCAGTCCATCTACATGGCCCAAGAGCCCGTGCA	563
Qy	662	GTTGCCAGACTATCACTTTGTGGACCATCGCATTTGAGATTGTGAGCCACGACAAAGATTA	721
Db	564	GCTCCCGGCTACTACTACGTGGACGCCAAGCTGGACATCACCTCCCAACACGAGGACTA	623
Qy	722	CAACAAGGTTAAGCTGCACGAGCATGCCGAAGCTCGTCATGGAAGT	768
Db	624	CACCATCGTGAGCAGTACGAGCGCACCGAGGGCGGCCACCACCTGT	670

Search completed: April 5, 2006, 21:43:39  
Job time : 669 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 20:50:59 ; Search time 4977 Seconds  
(without alignments)  
10021.103 Million cell updates/sec

Title: US-10-757-356-17

Perfect score: 1066

Sequence: 1 attcgccctgggatttggg.....atcgagcgcaaaaaaaa 1066

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	244.6	22.9	556	8	DR987866
2	241.4	22.6	556	8	DR987865
3	58	5.4	436	8	DR988183
4	47.4	4.4	961	10	CL209197
5	45.8	4.3	1201	10	CNS01078
6	45.4	4.3	1201	10	CNS01613
7	44.4	4.2	319	8	DR986899
8	44	4.1	678	8	DN886937
9	43.4	4.1	802	11	CNS0383B
10	43.2	4.1	895	9	CC154760
11	43.2	4.1	1101	10	CNS0039G
12	42.8	4.0	443	2	BF588024
13	42.8	4.0	480	2	BF587508
14	42.8	4.0	517	2	BF587613
15	42.8	4.0	692	2	BF586807
16	42.4	4.0	395	1	AJ651923
17	42.4	4.0	397	1	AJ651934
18	42.4	4.0	482	1	AJ658699
19	42.4	4.0	660	5	BQ604452
20	42.4	4.0	707	6	CF789251
21	42.2	4.0	772	10	CNS016XS
22	42	3.9	737	5	BQ955633

23	41.6	3.9	1101	10	CNS016K3
24	41.2	3.9	504	7	CV082013
C 25	41	3.8	725	10	AG183041
C 26	41	3.8	1101	10	CNS0021J
C 27	41	3.8	1101	10	CNS00ESI
C 28	40.6	3.8	1101	10	CNS016XR
C 29	40.6	3.8	1188	10	AG305240
C 30	40.6	3.8	1366	9	CC228527
C 31	40.4	3.8	1101	10	CNS0182P
C 32	40.2	3.8	387	6	CA651996
C 33	40.2	3.8	603	1	AW561384
C 34	40.2	3.8	644	3	BJ591738
C 35	40.2	3.8	711	3	BJ579615
C 36	40.2	3.8	743	3	BJ167081
C 37	40.2	3.8	928	10	CNS0102F
C 38	40	3.8	240	1	AV366735
C 39	40	3.8	578	5	C91535
C 40	40	3.8	595	10	AG231618
41	40	3.8	668	3	BJ345340
42	40	3.8	706	3	BJ375096
43	39.8	3.7	647	3	BM499773
44	39.8	3.7	736	10	CL582087
C 45	39.6	3.7	444	2	BI076362

#### ALIGNMENTS

RESULT 1  
DR987866  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

DR987866 556 bp mRNA linear EST 03-AUG-2005  
JGI\_AOSF1131.fwd AOSF Montastraea faveolata adult colony  
Montastraea faveolata cDNA clone AOSF1131 5', mRNA sequence.  
DR987866  
DR987866.1 GI:71782572  
EST.  
Montastraea faveolata  
Montastraea faveolata  
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;  
Favilina; Faviidae; Montastraea.  
1 (bases 1 to 556)  
Schwarz, J.A., Brokstein, P., Manohar, C., Coffroth, M.A., Szman, A.  
and Medina, M.  
Coral-Symbiodinium EST Project  
Unpublished (2005)  
Other ESTs: JGI\_AOSF1131.rev  
Contact: Schwarz, JA, Medina, M.  
Evolutionary Genomics  
DOE Joint Genome Institute  
2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
Tel: 925-296-5823  
Email: jaszwarz@lbl.gov  
cDNA Library Preparation: DOE Joint Genome Institute:  
<http://www.jgi.doe.gov>  
DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>  
Naming Conventions: EST name is generated by the concatenation of the JGI Clone id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.  
Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.  
Place: AOSF 0009 row: f column: 20  
High quality sequence stop: 557.  
Location/Qualifiers  
1. 556  
/organism="Montastraea faveolata"  
/mol\_type="mRNA"  
/db\_xref="taxon:48498"  
/clone="AOSF1131"  
/dev\_stage="Adult colony"  
/lab\_host="ElectroMAX DH10B"  
/clone\_lib="AOSF Montastraea faveolata adult colony"

/note=vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI; The library was prepared from total RNA using the Creator SMART cDNA Library Construction Kit with the LD-PCR method to amplify the cDNA. Amplified cDNA was digested with SfiI, size selected for >400bp, and ligated into the pDNR-LIB vector. WARNING: this library contains a small percentage of cDNAs derived from the coral's symbiont, Symbiodinium sp."

## ORIGIN

Query Match	22.9%	Score 244.6	DB 8	Length 556
Best Local Similarity	68.0%	Prod. No. 3.6e-58		
Matches 433	Conservative	0	Mismatches 114	Indels 90
				Gaps 3
QY	183	GATGGAAAAGGGAAGCCCTTTCCAGCGAACAACAGACTATGGACCTTACAGTCATAGAAGGC	242	
DB	1	GACGGAGNAGCCATTTTCGAGGGAACAACAGTATTAACCTTCAAGTCCAGAGGT	60	
QY	243	GCACCATTCCTTTTCGCTTACGATATCTTGTACAACAGTATTCGATTTACGGCAACAGGGTA	302	
DB	61	GGACCTCTCCCTTTTGTCTTACGATATCTTGAACGACAGCATTTCTCGTACGGCAACAGGGCA	120	
QY	303	TTGCGCAAAATACCCAGAGACATAGCAGATTAATTTCAAGCAGACGTTTCTCTGAGGGGTAC	362	
DB	121	TTACCAAAATACCCAGAGACCTCCAGACTTTTCAAGCAGTCTTTCTCGGGGGTAT	180	
QY	363	TTCTGGGAACGAACATGACATACGAAGACACAGGGCATTTTGCATCGGCCAACAACGACATA	422	
DB	181	TCCTGGGAACGTTGTATGACTTTTCGAAGATGGAGGCATTTTGCACCGTGTCAAAACAACATC	240	
QY	423	ACAAATGATGNAAGCGCTCGACGACTGTTTTCCTTATAAAATTCGATTTGATGGTGAAC	482	
DB	241	AAAATG-----GAGGGCGACCGTTTATCTATGATATTCGATTTTCATGGTGTAAAC	291	
QY	483	TTTCTCTGCCAATGGTCCAGTTTATGCAGAGGAAGACGCTGAAATGGGAGCCATCCACTGAG	542	
DB	292	TTTCTCTGCTAATGGTCCAAATTATCGAAGAAGA-----	325	
QY	543	ATAATGTATGGCGGTATGGAGTCTGAAGGTGATTTAACTGGCTCTGTTGCTTGAA	602	
DB	326	-----CTCTGTTGCTTTGAA	339	
QY	603	GGAGGTGGCCATTACCGATGTGACTTTCAAAACTATTACAAGCTAAGAAGTTTGTCCGG	662	
DB	340	GGAAATAAGCATCACCGATGTAACTTCAGAGTACTTACAAGCCAAGAAGGATGTGGTG	399	
QY	663	TTGCCAGACTATCATCTTTGTGGACCATCGCATTGAGATTGTGAGCCACGACAAGATTAC	722	
DB	400	TTGCCAAGATATCATCTTTGTTGACCCCGAATTGAGATATTAAGTCATGACAAAGATTAC	459	
QY	723	AACAAGGTTAAGCTGACGACATGCGGAAGCTCG-----TCATGGACTCTTCAAGG	773	
DB	460	AACAACTGGTGGTCTATGAAATGCTGCTGCCGCCCTCAGGCTTCTACTCTCCGAGC	519	
QY	774	AAGGCCAAGTAAAGGCTTAATGAAAAAGTCAAGACGAC	810	
DB	520	AAGGACAAGTAAAGGCTTAACGAAGAGCCAAACGAC	556	

## RESULT 2

Accession	DR987865	556 bp	mRNA	linear	EST 03-AUG-2005
Version	JGI AOSF1131.rev	AOSF	Montastraea	faveolata	adult colony
Keywords	Montastraea faveolata	cdna	clone	AOSF1131.3'	mRNA sequence.
Source	DR987865.1	GI:71782571			
Organism	Montastraea faveolata				
	Montastraea faveolata				
	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;				
	Favina; Faviidae; Montastraea.				
Reference	1 (bases 1 to 556)				
Authors	Schwarz, J.A., Brokstein, P., Manohar, C., Coffroth, M.A., Szmant, A.				

```
Db 231 -----CTCTGTGCTTGAA 218
QY 603 GGAGTGGCCATTACCGATGTGACTTCAAACTACTTACAAAGCTAAGAGGTGTCGCG 662
Db 217 GGAATAAGCATCACCAGTAACTTCACTAGAGTACTTACAAAGCTAAGAGGTGTCG 158
QY 663 TTGCCAGACTATCAGCTTTGTGGACCATCGCATTTGAGATTGTGAGCCACCAAGATTAC 722
Db 157 TTGCCAGATATCAGCTTTGTGGACCATCGCATTTGAGATTGAGTATTAAGTCATGACAAAGATTAC 98
QY 723 AACAGGTTAAGTGCAGAGCATGCGGAGCTCG-----TCATGAGCTGTCAAGG 773
Db 97 AACACGTGTGTCTATGAGATGTCTGCTGCCGCCCTCAGGCTTCTACTCTGCGGAGC 38
QY 774 AAGGCCAAGTAAAGGCTTAATGAAGAGTCAAGACGAC 810
Db 37 AAGGACAGTAAGGCTTACGAGAGGCCAACCCGCC 1

RESULT 3
DR988183
LOCUS JGI_AOSF1327.fwd AOSF Montastraea faveolata adult colony
DEFINITION Montastraea faveolata cDNA clone AOSF1327 5', mRNA sequence.
ACCESSION DR988183
VERSION DR988183.1 GI:71782889
KEYWORDS EST.
SOURCE Montastraea faveolata
ORGANISM Montastraea faveolata
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
Favini; Faviidae; Montastraea.
REFERENCE 1 (bases 1 to 436)
AUTHORS Schwarz, J.A., Brokstein, P., Manohar, C., Coffroth, M.A., Szman, A.
and Medina, M.
TITLE Coral-Symbiodinium EST Project
JOURNAL Unpublished (2005)
COMMENT Contact: Schwarz, J.A., Medina, M.
Evolutionary Genomics
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925-296-5823
Email: jaschwarz@lbl.gov
cDNA Library Preparation: DOE Joint Genome Institute:
http://www.jgi.doe.gov
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Small Insert: Based upon one or more sequencing reads of this clone
where vector sequence was present at both ends, this clone has been
determined to contain a cDNA insert on the order of 600-1000 bases.
Plate: AOSF 0013 row: n column: 19
High quality sequence stop: 436.
FEATURES
source
1. 436
/organism="Montastraea faveolata"
/mol_type="mRNA"
/db_xref="taxon:48498"
/clone="AOSF1327"
/dev_stage="Adult colony"
/lab_host="ElectroMAX DH10B"
/clone_lib="AOSF Montastraea faveolata adult colony"
/note="Vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI; The
library was prepared from total RNA using the Creator
SMART cDNA Library Construction Kit with the LD-PCR method
to amplify the cDNA. Amplified cDNA was digested with
SfiI, size selected for >400bp, and ligated into the
pDNR-LIB vector. WARNING: this library contains a small
percentage of cDNAs derived from the coral's symbiont,
Symbiodinium sp."
```

```
Query Match 5.4%; Score 58; DB 8; Length 436;
Best Local Similarity 65.0%; Pred. No. 5.7e-05;
Matches 143; Conservative 0; Mismatches 55; Indels 22; Gaps 3;

QY 765 CTGTCAAGGAAGCCCAAGTAAAGGCTTA-ATGAAAAGTCAAGACGACGAGGAGAAAC 823
Db 32 CTGCCAAGCAAGCCAAAGTAAAGGCTTAGCGCAATAGTCAAGACGACCAACAAGAGAAA 91
QY 824 AAAGTACTTTTTTTTAAATTTGAAGGCATTTACTCGGAATTAGTATTGTACTTTTGA 883
Db 92 TAAAAGTTGTCAGAGTTAGAAATTTGATATTCAGT-----TTTCGA 132
QY 884 TTCAAGAGTTTGTCCCGGATTTGTAGAGACTAGCTCTAGAGTTGTATTGTGAAAAA 943
Db 133 TTCAAGGATTCGTTACGGGTTTGTGTAGAGACTAGCTTTAGAGT--TAAGTTTGCAGAA 190
QY 944 AGATAGTTCCAGTTTTTTCGGGATTTACAGCATGGGGATA 983
Db 191 AAAAGGTTTCCATTTTTCGCCGCTTTATAGCATAAGAGTA 230

RESULT 4
CL209197/c
LOCUS ZM988183.1 ZM988183 (HindIII) Zea mays genomic clone
DEFINITION ZM988183.1 3', genomic survey sequence.
ACCESSION CL209197
VERSION CL209197.1 GI:40725634
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 961)
AUTHORS Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
Zohovetz, V., Fuke, G., Yu, Y., Wing, R. and Messing, J.
TITLE Sequencing of the maize genome at PGIR (2003c)
JOURNAL Unpublished (2003)
COMMENT Contact: Bharti, A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 101.
FEATURES
source
1. 961
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZM988183.1"
/lab_host="E. coli DH10B"
/clone_lib="ZM988183 (HindIII)"
/note="Vector: pCUGI; Site 1: HindIII; Site 2: HindIII"

ORIGIN
Query Match 4.4%; Score 47.4; DB 10; Length 961;
Best Local Similarity 53.3%; Pred. No. 0.077;
Matches 122; Conservative 0; Mismatches 106; Indels 1; Gaps 1;

QY 815 AGGAGAAACAAAGTACTTTTTTTTAAATTTGAAGGCATTTACTCGGAATTAGTATTGA 874
Db 599 ATGTCATCTAAGAGGATGTCCTTTTGAATTCATTTTCTGATAATAGATGA 540
QY 875 TACTTTTCGATTCAAGGATTTTTCGGGATTTGTAGAGACTAGCTCTAGAGTTGTATT 934
Db 539 TAATATTAAATTTTGAATGTTCAAGTTTAGATTGTAATATTGGCTTTAGAGTTGAATT 480
```



```

Qy 935 TGTGAAAAAGATAGTTTCCAGTTTTCGGGATTACAGCATGGGATAGACTTTT-TAA 993
Db 479 TAAGTAATAGAAAGATTTTCATTTATTGGAGTGATTTATTATATATATATAT 420

Qy 994 ACTCAGTTGTGTCGTAATGCAAGTGAAGAAACCTGTAGTGAGATAAAT 1042
Db 419 AATTAGAGTATTTATATACAAAGTGAAGAAACAAATAAATAGATGAAT 371

RESULT 5
CNS0107R/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN03G13 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL098625.1 GI:5610236
VERSION
AL098625
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1201)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
Location/Qualifiers
source
1..1201
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN03G13"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

ORIGIN
Query Match 4.3%; Score 45.8; DB 10; Length 1201;
Best Local Similarity 30.9%; Pred. No. 0.23; Mismatches 87; Indels 6; Gaps 1;
Matches 77; Conservative 79; Mismatches 87; Indels 6; Gaps 1;

Qy 798 AAGTCAAGCAGCAACGAGGAGAAACAAAGTACTTTTGTGTAATTTGAAAGCATTTAC 857
Db 998 WDDWTWARGWRAVAAARRRGVGSARDDVAKTTTDDVDDTTTWTWAAWABAR 939

Qy 858 TCGGAATTAGTATTGATATTCGATTCAGGATTGTTCCGGGATTTGTAGAGACTA 917
Db 938 WWDATTTTDTATTTGATATTTTAAKTRRDARDAADWKTTTGTGGDTTKAKTWTGRGWTW 879

Qy 918 -----GCTCTAGAGTGTATTTCGTGAAAAAGATAGTTTCCAGTTTTCGGGATTAC 971
Db 878 ATGGGGRGDMWTGGVKGGGDTTTRRSAGADWAWAAWTTTGATWATTTTARTTTAA 819

Qy 972 AGCATGGGGATAGACTTTTAACTCAGTTGGTCAATGCAAGTGAAGAAACCTGTAGT 1031
Db 818 ADTTRRRAGWGTDDTGKTEGAGGGWAGGDDWAAATRGGAWAAAKTAAATTTATKBK 759

Qy 1032 GAGATATA 1040
Db 758 KGAATAARA 750

```

```

RESULT 6
CNS01613/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN15108 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL106161
VERSION
AL106161.1 GI:5620457
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1201)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
Location/Qualifiers
source
1..1201
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15108"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

ORIGIN
Query Match 4.3%; Score 45.4; DB 10; Length 1201;
Best Local Similarity 39.3%; Pred. No. 0.31; Mismatches 35; Indels 0; Gaps 0;
Matches 68; Conservative 35; Mismatches 70; Indels 0; Gaps 0;

Qy 780 AAGTAAAGGCTTAATGAAAAGTCAAGCAGCAGAGAGAAACAAAGTACTTTTGT 839
Db 1162 AAATDGGGAAAAAAAATAATGTRAAAAATTTGAATAATGAAWAAAAAATTKT 1103

Qy 840 AAATTTGAAGGCATTTTACTCGGAATTAGTATTGATATTCGATTCAAGGATTTGTC 899
Db 1102 WAATKCAAWWAWWTWDTTWDAAWAKATAATTTTWTWTTTAAWTKATGKKTKTK 1043

Qy 900 GGGATTTGTAGAGACTAGCTCTAGACTGTATTGTTGTAAGAAAAAGATAGTT 952
Db 1042 TTTTKTITTTTKRGKTKTTTMAATTTTCTTGTGKDAWAAAAAAGAGT 990

RESULT 7
DR986899
LOCUS
DEFINITION
JGI AOSC1075.fwd AOSC Montastraea faveolata 60 hours post
fertilization Montastraea faveolata cDNA clone AOSC1075 5', mRNA
sequence.
ACCESSION
DR986899.1 GI:71781605
VERSION
DR986899
KEYWORDS
EST.
SOURCE
Montastraea faveolata
ORGANISM
Montastraea faveolata
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Favina; Faviidae; Montastraea.
REFERENCE
1 (bases 1 to 319)
AUTHORS
Schwarz, J.A., Brokstein, P., Manohar, C., Coffroth, M.A., Szmant, A.
and Medina, M.

```



Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE  
AUTHORS  
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzanes, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W., and Weissenbach, J.  
TITLE  
Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence  
JOURNAL  
Nat. Genet. 25 (2), 235-238 (2000)  
PUBMED  
10835645

REFERENCE  
AUTHORS  
Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fzanes, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.  
TITLE  
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
JOURNAL  
Genome Res. 10 (7), 939-949 (2000)  
PUBMED  
10899143

REFERENCE  
AUTHORS  
3 (bases 1 to 802)  
Genoscope.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT  
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.  
FEATURES  
source  
Location/Qualifiers  
1..802  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="004012"  
/clone\_lib="G"  
/note="Genoscope sequence ID : CORG004BH06XE1 end : PUC-ori"

ORIGIN  
Query Match 4.1%; Score 43.4; DB 11; Length 802;  
Best Local Similarity 41.9%; Pred. No. 1;  
Matches 101; Conservative 22; Mismatches 118; Indels 0; Gaps 0;  
QY 825 AAGTACTTTTGTGTTAAATTTGAAGGCATTACTCGGAATTTAGTATTTGATCTTCGAT 884  
DB 320 AAKTTTATTTTATTTTATTTTATGKWKIGAAKTWATAAAAAATAWAAGATTTTAT 261  
QY 885 TCAAGGATTGTTTCGGGATTGTTAGAGACTAGCTCTAGAGTTGTATTTTGTGAAAAA 944  
DB 260 TAAAGTKATTTTTTTTATTTTATTTTATTTTAAAGKATTTTAAAGKTTTAATA 201  
QY 945 GATAGTTTCCAGTTTTCGGGATTACAGCATGGGATAGACTTTTAAACTCAGTTGTG 1004  
DB 200 AATTTTCTTTGTTGGAGAGGTGTGTAGTGGGGGGTGAATGTGTAGGGTRARATGTG 141  
QY 1005 GTCAAATCCAGTAAAGAACTGTAGTCGATAAATCTGTTATCGAGCCGAAAAAAA 1064  
DB 140 RAGAAARARAGGAAAAAAGAAAGAAAGAGRAGTTKTTTAAAAAGRGTTAATAT 81  
QY 1065 A 1065  
DB 80 A 80

RESULT 10  
CC154760/c  
LOCUS  
DEFINITION  
CC154760 895 bp DNA linear GSS 25-APR-2003  
CSU-K34.122B16.SP6 CSU-K34 Aedes aegypti genomic clone  
CSU-K34-122B16, genomic survey sequence.  
ACCESSION  
CC154760  
VERSION  
CC154760.1 GI:30108056  
KEYWORDS  
GSS.  
SOURCE  
Aedes aegypti (yellow fever mosquito)  
ORGANISM  
Aedes aegypti

Tetradontidae; Tetraodontidae; Tetraodon.

REFERENCE  
AUTHORS  
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzanes, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W., and Weissenbach, J.  
TITLE  
Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence  
JOURNAL  
Nat. Genet. 25 (2), 235-238 (2000)  
PUBMED  
10835645

REFERENCE  
AUTHORS  
Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fzanes, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.  
TITLE  
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
JOURNAL  
Genome Res. 10 (7), 939-949 (2000)  
PUBMED  
10899143

REFERENCE  
AUTHORS  
3 (bases 1 to 802)  
Genoscope.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT  
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.  
FEATURES  
source  
Location/Qualifiers  
1..802  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="004012"  
/clone\_lib="G"  
/note="Genoscope sequence ID : CORG004BH06XE1 end : PUC-ori"

ORIGIN  
Query Match 4.1%; Score 43.4; DB 11; Length 802;  
Best Local Similarity 41.9%; Pred. No. 1;  
Matches 101; Conservative 22; Mismatches 118; Indels 0; Gaps 0;  
QY 825 AAGTACTTTTGTGTTAAATTTGAAGGCATTACTCGGAATTTAGTATTTGATCTTCGAT 884  
DB 320 AAKTTTATTTTATTTTATTTTATGKWKIGAAKTWATAAAAAATAWAAGATTTTAT 261  
QY 885 TCAAGGATTGTTTCGGGATTGTTAGAGACTAGCTCTAGAGTTGTATTTTGTGAAAAA 944  
DB 260 TAAAGTKATTTTTTTTATTTTATTTTATTTTAAAGKATTTTAAAGKTTTAATA 201  
QY 945 GATAGTTTCCAGTTTTCGGGATTACAGCATGGGATAGACTTTTAAACTCAGTTGTG 1004  
DB 200 AATTTTCTTTGTTGGAGAGGTGTGTAGTGGGGGGTGAATGTGTAGGGTRARATGTG 141  
QY 1005 GTCAAATCCAGTAAAGAACTGTAGTCGATAAATCTGTTATCGAGCCGAAAAAAA 1064  
DB 140 RAGAAARARAGGAAAAAAGAAAGAAAGAGRAGTTKTTTAAAAAGRGTTAATAT 81  
QY 1065 A 1065  
DB 80 A 80

RESULT 10  
CC154760/c  
LOCUS  
DEFINITION  
CC154760 895 bp DNA linear GSS 25-APR-2003  
CSU-K34.122B16.SP6 CSU-K34 Aedes aegypti genomic clone  
CSU-K34-122B16, genomic survey sequence.  
ACCESSION  
CC154760  
VERSION  
CC154760.1 GI:30108056  
KEYWORDS  
GSS.  
SOURCE  
Aedes aegypti (yellow fever mosquito)  
ORGANISM  
Aedes aegypti

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae; Culicini; Aedes; Stegomyia.

REFERENCE  
AUTHORS  
Loftus, B., Shetty, J., Knudson, D. and Severson, D.  
TITLE  
BAC end sequencing of Aedes aegypti  
JOURNAL  
Unpublished (2003)  
COMMENT  
Other GSSs: CSU-K34.122B16.T7  
Contact: Brendan Loftus  
Department of Eukaryotic Genomics  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: enta@tigr.org  
Library was provided by Susan Brown and Dennis Knudson at Colorado State University.  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
Location/Qualifiers  
1..895  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7159"  
/clone="CSU-K34-122B16"  
/clone\_lib="CSU-K34"  
/note="Vector: pBACe3.6; Site 1: EcoRI; Source DNA: Aedes aegypti; strain unknown [derived from freshly hatched larvae at the Virus Research Centre, Poona, India. Reference: SINGH, K. R. P., 1967 Cell cultures derived from larvae of Aedes albopictus (Skuse) and Aedes aegypti (L.). Current Science 36: 506-508]; ATC-10 cell line ATCC CCL-125"

ORIGIN  
Query Match 4.1%; Score 43.2; DB 9; Length 895;  
Best Local Similarity 48.4%; Pred. No. 1.2;  
Matches 120; Conservative 0; Mismatches 128; Indels 0; Gaps 0;  
QY 818 AGAACAAAGTACTTTTGTGTTAAATTTGAAGGCATTACTCGGAATTTAGTATTTGATAC 877  
DB 367 ATAATTAATTAATAATGTTTGCATTTCAATTCGAAAAAGTTCTAAATTAGTCATGTTGT 308  
QY 878 TTTTCGATTCAAGATTGTTTCCGGGATTGTTAGAGACTAGCTCTAGAGTTGATTTTGT 937  
DB 307 TACAAATTTGCAATTTAAAGTTGTTCAAAATTCAGTTTCCATTTTCTTTTCT 248  
QY 938 GAAAAAGATAGTTTCCAGTTTTCGGGATTACAGCATCGGGATAGACTTTTAAACTC 997  
DB 247 GGAATTAATTTTCCCAATAATTTTGAATAATTTGTCATGAAGGCTTTCACCATCATGAG 188  
QY 998 AGTTGTGTCGAATGCAAGTAAGAAACTGTAGTGAAGTAAGAACTTTTATCGAGCCGA 1057  
DB 187 TTTCTTCAAAATACGAAATTTAGAACTCTTTCCAGAAAAAAGTGAATTAAGAAAGAAC 128  
QY 1058 AAAAAAA 1065  
DB 127 CAAAAAAA 120

RESULT 11  
CNS0039G/c  
LOCUS  
DEFINITION  
CNS0039G 1101 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TTT3 end of BAC # BACR08K10 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION  
AL063921  
VERSION  
AL063921.1 GI:4941778  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;



LOCUS BF587508 480 bp mRNA linear EST 12-DEC-2000  
 DEFINITION F01\_37\_C08.g1 A003 Floral-Induced Meristem 1 (F01) Sorghum  
 propinquum cDNA, mRNA sequence.  
 ACCESSION BF587508  
 VERSION BF587508.1 GI:11679832  
 KEYWORDS EST.  
 SOURCE Sorghum propinquum  
 ORGANISM Sorghum propinquum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 480)  
 AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M., Marsala, C. and  
 Pratt, L.H.  
 TITLE An EST database from Sorghum: floral-induced meristems  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: PolyTWix  
 High quality sequence start: 5  
 High quality sequence stop: 477  
 POLYA=No.

FEATURES  
 source  
 1..480 Location/Qualifiers  
 /organism="Sorghum propinquum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:132711"  
 /note="Organ: Floral-Induced Meristem 1 (F01)"  
 pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2:  
 EcoRI; mature plants were placed in a growth chamber for  
 15 days with 16 hr darkness, and 8 hr light (flowering is  
 induced by short-day conditions); 16 days after being  
 returned to the greenhouse under natural long days during  
 late April/early May, meristems were harvested. The  
 library was made from poly-A RNA in the cloning vector  
 lambda ZAP II. Clones to be sequenced were prepared by  
 mass excision."

ORIGIN  
 Query Match 4.0%; Score 42.8; DB 2; Length 480;  
 Best Local Similarity 54.4%; Pred. No. 1.3;  
 Matches 86; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
 Qy 55 TTGATATCTTACTTACGCTACCATCATGACAAAGTGTTCACAGGAAAGGGTGTGATTA 114  
 Db 294 TAGATAACCTAGTACCTTATGCTTGGTACGAGCTAGCACTCGTGAAGAGGGTGCTTA 235  
 Qy 115 AACACAGATGAGATGAGCTGCTATGGAAGTGTGTAAACGGGCACAAAGTTCGTGG 174  
 Db 234 GGTGAGTCGATCACAGTTTACAGCTCCGGCTGGGTGCTATAGGCTTCAGGAAGTTA 175  
 Qy 175 TTGAAGGAGATGGAAGAGGAGCGCTTTTCGACGGAACA 212  
 Db 174 TCGAAGGGGCTGGCAGGTGGAGCTTTTACTACAGNACA 137

RESULT 14  
 BF587613/c  
 LOCUS BF587613 517 bp mRNA linear EST 12-DEC-2000  
 DEFINITION F01\_38\_F03.g1 A003 Floral-Induced Meristem 1 (F01) Sorghum  
 propinquum cDNA, mRNA sequence.  
 ACCESSION BF587613  
 VERSION BF587613.1 GI:11679937  
 KEYWORDS EST.

Sorghum propinquum  
 Sorghum propinquum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 517)  
 AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M., Marsala, C. and  
 Pratt, L.H.  
 TITLE An EST database from Sorghum: floral-induced meristems  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: PolyTWix  
 High quality sequence start: 132  
 High quality sequence stop: 511  
 POLYA=No.

FEATURES  
 source  
 1..517 Location/Qualifiers  
 /organism="Sorghum propinquum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:132711"  
 /note="Organ: Floral-Induced Meristem 1 (F01)"  
 pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2:  
 EcoRI; mature plants were placed in a growth chamber for  
 15 days with 16 hr darkness, and 8 hr light (flowering is  
 induced by short-day conditions); 16 days after being  
 returned to the greenhouse under natural long days during  
 late April/early May, meristems were harvested. The  
 library was made from poly-A RNA in the cloning vector  
 lambda ZAP II. Clones to be sequenced were prepared by  
 mass excision."

ORIGIN  
 Query Match 4.0%; Score 42.8; DB 2; Length 517;  
 Best Local Similarity 54.4%; Pred. No. 1.3;  
 Matches 86; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
 Qy 55 TTGATATCTTACTTACGCTACCATCATGACAAAGTGTTCACAGGAAAGGGTGTGATTA 114  
 Db 396 TAGATAACCTAGTACCTTATGCTTGGTACGAGCTAGCACTCGTGAAGAGGGTGCTTA 337  
 Qy 115 AACACAGATGAGATGAGCTGCTATGGAAGTGTGTAAACGGGCACAAAGTTCGTGG 174  
 Db 336 GGTGAGTCGATCACAGTTTACAGCTCCGGCTGGGTGCTATAGGCTTCAGGAAGTTA 277  
 Qy 175 TTGAAGGAGATGGAAGAGGAGCGCTTTTCGACGGAACA 212  
 Db 276 TCGAAGGGGCTGGCAGGTGGAGCTTTTACTACAGNACA 239

RESULT 15  
 BF586807/c  
 LOCUS BF586807 692 bp mRNA linear EST 12-DEC-2000  
 DEFINITION F01\_30\_E11.g1 A003 Floral-Induced Meristem 1 (F01) Sorghum  
 propinquum cDNA, mRNA sequence.  
 ACCESSION BF586807  
 VERSION BF586807.1 GI:11679033  
 KEYWORDS EST.  
 SOURCE Sorghum propinquum  
 ORGANISM Sorghum propinquum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 692)

## AUTHORS

Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M., Marsala, C. and Pratt, L.H.  
An EST database from Sorghum: floral-induced meristems  
Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: PolyTMix  
High quality sequence start: 62  
High quality sequence stop: 691  
POLYA=No.

## FEATURES

## source

## Location/Qualifiers

1..692  
/organism="Sorghum propinquum"  
/mol\_type="mRNA"  
/db\_xref="taxon:132711"  
/clone\_lib="Floral-Induced Meristem 1 (FM1)"  
/note="Organ: Floral-induced meristems; Vector: pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; mature plants were placed in a growth chamber for 15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May, meristems were harvested. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

## ORIGIN

Query Match 4.0%; Score 42.8; DB 2; Length 692;  
Best Local Similarity 54.4%; Pred. No. 1.4;  
Matches 86; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 55 TTGATATCTTACTTACGTTACCATCATGACAAGTGTTCACAGGAAAAGGGTGTGATTA 114  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
536 TAGATAACCTAGTACCCTATGCTTTGGTACGCACTCTAGCACTCGTGAAGAGGGTGCTTA 477  
QY 115 AACGACATGACATGAAGTCGCTATGGAGTGTCTTAACGGGCACAAAGTTCGTGG 174  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
476 GGTCAAGTCGATCACAGTTCACAGTCCCGGCTGGTGTCTATAGGCTTCAGGAAGTTA 417  
QY 175 TTGAAGGAGATGGAAGGGAAGCCTTTTCGACGGAACA 212  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
416 TCGAAGGGGCTGGCAGGTGGNAGCTTTACTACAGAACA 379

Search completed: April 5, 2006, 22:24:01  
Job time : 4982 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 20:42:51 ; Search time 3829 Seconds

(without alignments)  
15825.312 Million cell updates/sec

Title: US-10-757-356-17

Perfect score: 1066

Sequence: 1 attcgccctgggtgattgga.....atcggaagcgcaaaaaaa 1066

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_scs.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1066	100.0	1066	2	AY037769	Montastraea
2	957.6	89.8	972	2	AY181554	Montastraea
3	948.6	89.0	1045	2	AF406766	Montastraea
4	914.8	85.8	1062	2	AY679109	Montastraea
5	876	82.2	972	2	AY181555	Montastraea
6	836.4	78.5	913	2	AY037768	Montastraea
7	834.2	78.3	1046	2	AF401282	Montastraea
8	681.2	63.9	729	11	AY182015	Synthetic
9	679.6	63.8	1045	2	AY181552	Montastraea
10	678.4	63.6	836	2	AY181553	Montastraea
11	665	62.4	1133	2	AY037766	Montastraea
12	625.4	58.7	1046	2	AF384683	Montastraea
13	624.8	58.6	684	2	AF679112	Montastraea
14	623.2	58.5	684	2	AY679110	Montastraea
15	621.8	58.3	708	11	AY182016	Synthetic
16	621.8	58.3	708	11	AY182021	Synthetic
17	610	57.2	967	2	AY181557	Montastraea
18	604.2	56.7	708	11	AY182014	Synthetic

19	541.6	50.8	684	2	AB193294	AB193294 Pavia fav
20	533.4	50.0	795	2	AY037770	AY037770 Montastraea
21	527.2	48.5	684	2	AB193293	AB193293 Pavia fav
22	521.8	48.9	879	2	AY181556	AY181556 Montastraea
23	521.6	48.9	675	2	AB180726	AB180726 Echinophy
24	513.6	48.2	702	11	AY182020	AY182020 Synthetic
25	507.4	47.6	729	6	CS079401	CS079401 Sequence
26	507.2	47.6	678	6	CS079393	CS079393 Sequence
27	507	47.6	678	2	AY679111	AY679111 Montastraea
28	501	47.0	705	11	AY181558	AY181558 Synthetic
29	499.2	46.8	702	11	AY182013	AY182013 Synthetic
30	497.8	46.7	705	11	AY182018	AY182018 Synthetic
31	497.8	46.7	729	6	CS079402	CS079402 Sequence
32	497.4	46.7	678	6	CS079394	CS079394 Sequence
33	489.2	45.9	702	11	AY182012	AY182012 Synthetic
34	488	45.8	675	2	AY362545	AY362545 Montastraea
35	486.4	45.6	678	2	AB107915	AB107915 Galaxea f
36	483.2	45.3	702	11	AY182019	AY182019 Synthetic
37	472	44.3	678	2	AB108447	AB108447 Galaxea f
38	469	44.0	705	11	AY182023	AY182023 Synthetic
39	467.2	43.8	675	6	CS079392	CS079392 Sequence
40	467.2	43.8	726	6	CS079400	CS079400 Sequence
41	464	43.5	705	11	AY182017	AY182017 Synthetic
42	459.2	43.1	675	2	AY056460	AY056460 Montastraea
43	456.2	42.8	705	11	AY182022	AY182022 Synthetic
44	454	42.6	681	2	AY765217	AY765217 Lobophyll
45	437.4	41.0	5070	11	AY585853	AY585853 Psilostrike

## ALIGNMENTS

RESULT 1  
AY037769  
LOCUS Montastraea cavernosa mcavFP\_6 mRNA, complete cds.  
DEFINITION Montastraea cavernosa mcavFP\_6 mRNA, complete cds.  
ACCESSION AY037769  
VERSION AY037769.1 GI:19982568  
KEYWORDS Montastraea cavernosa (great star coral)  
SOURCE Montastraea cavernosa (great star coral)  
ORGANISM Montastraea cavernosa  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia; Faviina; Faviidae; Montastraea.  
REFERENCE 1 (bases 1 to 1066)  
AUTHORS Labas Y.A., Gurskaya, N.G., Yanushevich, Y.G., Pradkov, A.F., Lukyanov, K.A., Lukyanov, S.A. and Matz, M.V.  
TITLE Diversity and evolution of the green fluorescent protein family  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4256-4261 (2002)  
PUBMED 11929996  
REFERENCE 2 (bases 1 to 1066)  
AUTHORS Matz, M.V. and Lukyanov, S.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Institute of Bioorganic Chemistry RAS, Miklukho-Maklaya 16/10, Moscow 117871, Russia

## FEATURES

source  
1. 1066  
/organism="Montastraea cavernosa"  
/mol\_type="mRNA"  
/db\_xref="taxon:63558"  
81..785  
/note="GFP-like protein; green"  
/codon\_start=1  
/product="mcavFP\_6"  
/protein\_id="AAK71335.1"  
/translation="MTSAQKGVKPKMKLMRMEGVANGHKFPVEGDKGKPGDGT  
QMDGLTVEGAPLFDITVTVDYGNRVFAKYPEDYADYFKOTFPFGYFWSMTY  
EDQICATIDITWMEGVDDCFYKIRFDGVNFPANGVMQRTKQKSTEIYARD  
GVLKGVNMLLLEGGGHCYRCDFTTYKAKKVVLPDYHFDHRIEIVSHDKDYNKVK  
LHEAERHGLSRKAK"

## ORIGIN

Query Match 100.0%; Score 1066; DB 2; Length 1066;





```

361 ACTTCTGGGAACGAAGCATGACATACGAAGACCAAGGCGATTTCATGCCACAAACGACA 420
QY
421 TAAACAATGATGGAAGGCGTCCAGCAGCTGTTTGGCTATATAAAATTCGATTTGATGGTGA 480
Db
421 TAAACAATGATGGAAGGCGTCCAGCAGCTGTTTGGCTATATAAAATTCGATTTGATGGTGA 480
QY
481 ACTTCTCTGCCAATGGTCCAGTATGACAGAGGAAGCGCTGAAATGGGAGCCATCCACTG 540
Db
481 ACTTCTCTGCCAATGGTCCAGTATGACAGAGGAAGCGCTGAAATGGGAGCCATCCACTG 540
QY
541 AGATAATGATGACCGGTGATGAGTGTGGAAGGCGTATGTTAAATGCGCTCTGTTGCTTG 600
Db
541 AGATAATGATGACCGGTGATGAGTGTGGAAGGCGTATGTTAAATGCGCTCTGTTGCTTG 600
QY
601 AAGGAGTGGCCATTACCGATGCTCAAACTACTTACAAAGCTTAAGAAGGTTGTCC 660
Db
601 AAGGAGTGGCCATTACCGATGCTCAAACTACTTACAAAGCTTAAGAAGGTTGTCC 660
QY
661 GGTGTCAGACTACTACTTGTGGACCATCGCATGAGATTGTGAGCCACGACAAAGATT 720
Db
661 GGTGTCAGACTACTACTTGTGGACCATCGCATGAGATTGTGAGCCACGACAAAGATT 720
QY
721 ACAACAAGGTTAAGTCCACGAGCATCCGAGCTCGTCAAGGAGGCGCA 780
Db
721 ACAACAAGGTTAAGTCCACGAGCATCCGAGCTCGTCAAGGAGGCGCA 780
QY
781 AGTAAGGCTTAATGAAAAGTCAAGACGACACGAGGAGAAAACAAAGTACTTTTGTGA 840
Db
781 AGTAAGGCTTAATGAAAAGTCAAGACGACACGAGGAGAAAACAAAGTACTTTTGTGA 840
QY
841 AATTGGAAGCATTTACTCGGAATTAGTATTGATCTTTCGATCAAGGAGTTTTCG 900
Db
841 AATTGGAAGCATTTACTCGGAATTAGTATTGATCTTTCGATCAAGGAGTTTTCG 900
QY
901 GGAATTTGTTAGAGTACTCTAGAGTTGATTTTGTGAAAAAGATGTTTCCAGTTT 960
Db
901 GGAATTTGTTAGAGTACTCTAGAGTTGATTTTGTGAAAAAGATGTTTCCAGTTT 960
QY
961 TGGCGGATTACA 972
Db
961 TGGCGGATTACA 972

RESULT 3
AF406766
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .1045
/organism="Montastraea cavernosa"
/mol_type="mRNA"
/db_xref="taxon:63558"
/notes="isolated from specimen with red color"
84. .767
CDS

```

```

/codon_start=1
/product="green fluorescent protein"
/protein_id="AAK97633.1"
/db_xref="GI:15425965"
/translation="MSVIRKPMKIKLRMEGAVNGHFVIEGDKGPKFFECTQSMDLTV
IEGALPFPAYDILITLVFYKIRFDGVNFPANGPMORKTLLKWEPSSTEIMYARDGVKGDV
ATNDITMMKGVDDCFVYKIRFDGVNFPANGPMORKTLLKWEPSSTEIMYARDGVKGDV
NNALLLEGSGHYRCDPKTYKAKKVVRLPDIYHFDVHRIEIVSHDKYNKVKLYEHAEA
XXGLSRKAK"
ORIGIN
Query Match      89.0%;   Score 948.6;   DB 2;   Length 1045;
Best Local Similarity 95.2%;   Pred. No. 2 5e-256;
Matches 104;   Conservative 4;   Mismatches 23;   Indels 24;   Gaps 3;
QY 1 ATTGCGCCCTGGTGATTTGGAAGAGAGCAGATCGAGAACAAACAGAGCTG--TAAAGTTGA 58
Db 2 ATTGCGCCCTGGTGATTTGGAAGAGAGCAGATCAAGAAACAAACAGAGCTGTATTAAGGTTTA 61
QY 59 TATCTTACTTACGCTCTACCATCATGACAAAGTGTTCACAGGAAAGGGTGTGATTAAC 118
Db 62 TATCTTACTTACGCTCTACCGTCATGA-----GTGTGATAAACC 100
QY 119 AGACATGAAGATGAAGCTCGGTATGGAAGGTGCTGTAAACGGGCAACAAGTTTCGTGTTGA 178
Db 101 AGACATGAAGATCAAGCTCGGTATGGAAGGCGCTGTAAACGGGCAACAAGTTTCGTGTTGA 160
QY 179 AGGAGATGAAAAGGGAAGCCCTTTGACGGAACACAGACTATGGACCTTACAGTCATAGA 238
Db 161 AGGAGACGGAAGGCGCAAGCCCTTTGAGGGAACACAGAGTATGGACCTTACAGTCATAGA 220
QY 239 AGCGCACCATTTGCCCTTTCGCTTACGATATCTTGACAAAGTATTCGATTACGGCAACAG 298
Db 221 AGCGCACCATTTGCCCTTTCGCTTACGATATCTTGACAAAGTATTCGATTACGGCAACAG 280
QY 299 GGTATTTGCGCAATATACCCAGAAAGACATAGCAGATTATTTCAAGCAGACAGCTTCTCGAGG 358
Db 281 GGTATTTGCGCAATATACCCAGAAAGACATACAGATTATTTCAAGCAGACAGCTTCTCGAGG 340
QY 359 GTACTTCTGGGAACGAAGCATGACATACGAAGACGAGGCAATTTGCATCGCCACAAACGA 418
Db 341 GTACTTCTGGGAACGAAGCATGACATACGAAGACGAGGCAATTTGCATCGCCACAAACGA 400
QY 419 CATACATATGATGGAAGGCGTCGACCTGTTTTCGCTATAAAATTCGATTTCGATGCTGT 478
Db 401 CATACATATGATGGAAGGCGTCGACCTGTTTTCGCTATAAAATTCGATTTCGATGCTGT 460
QY 479 GAACTTTCTGCAATGGTCCAGTTTATGAGAGGAGAGCGCTGAAATGGGAGCGCATCCAC 538
Db 461 GAACTTTCTGCAATGGTCCAGTTTATGAGAGGAGAGCGCTGAAATGGGAGCGCATCCAC 520
QY 539 TGAGATAATGATATGCGCGTGATGGAGTGTGTAAGGGTGATGTTAAACATGGCTCTGTGTCT 598
Db 521 TGAGATAATGATATGCGCGTGATGGAGTGTGTAAGGGTGATGTTAAACATGGCTCTGTGTCT 580
QY 599 TGAAGAGGTGGCCATTACCGATGACCTTCAAACTACTTACAAAGCTTAAGAGGTTGT 658
Db 581 TGAAGAGGTGGCCATTACCGATGACCTTCAAACTACTTCAAAAGCTTAAGAGGTTGT 640
QY 659 CGGTTTGCAGACTATCACTTTGTGGACCATCGCATTTGAGATTGAGAGCCACGACAAAGA 718
Db 641 CGGTTTGCAGACTATCACTTTGTGGACCATCGCATTTGAGATTGAGAGCCACGACAAAGA 700
QY 719 TTAACAACAGGTTAAGTGTGACGAGCATGCCGAAGCTCGTCAATGGAAGTCTCAAGGAAGC 778
Db 701 TTAACAACAGGTTAAGTGTGATGAGCATGCCGAAGCTCTCTGTCGACTGTCTCAGGAAGC 760
QY 779 CAAGTAAAGGCTTAATGAAAAGTCAAGACGACGAGGAGAAACA-ANGTACTTTTGT 837
Db 761 CAAGTAAAGGCTTAATGAAAAGTCAAGACGACGAGGAGAAACAACAAGTAGTGT 820
QY 838 TTAATTTGAAGGCAATTTACTCGGAATAGTATTGATGTTTTCGATTCAAGGATTTGT 897

```

```

Db      821  TTAATTTGAAGGCATTACTCGCAATTAGTATTATTGATATCTTCGATTCAAGGATTGTT 880
Qy      898  CCGGATTTGTAGAGACTAGCTCTAGAGTTGATTTTGTGCAAAAAGATAGTTCCAGT 957
Db      881  TCGGGATTTGTAGAGACTAGCTCTAGAGTTGATTTTGTGCAAAAAGATAGTTCCAGT 940
Qy      958  TTTTGGCGGATTTACAGCATGGGATAGACTTTTAAACTCAGTTGTGCTCAAAATGCAAGT 1017
Db      941  TTTTGGCGGATTTACAGCATGGGATAGACTTTTAAACTCAGTTGTGCTCAAAATGCAAGT 1000
Qy      1018  AAGAAAATCTGTAGTGAAGTAACTTTGTTATCGAAGCGCGAAAAA 1062
Db      1001  AAGAAAATCTGTAGTGAAGTAACTTTGTTATCGAAGCGCGAAAAA 1045

RESULT 4
AY679109 1062 bp mRNA linear INV 25-AUG-2004
LOCUS Montastraea cavernosa green fluorescent protein G1 mRNA, complete cds.
DEFINITION
ACCESSION AY679109
VERSION AY679109.1 GI:51472050
KEYWORDS
SOURCE Montastraea cavernosa (great star coral)
ORGANISM Montastraea cavernosa
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Faviina; Faviidae; Montastraea.
REFERENCE
1 (bases 1 to 1062)
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.
TITLE Cloning of Cnidarian Fluorescent Protein Genes
JOURNAL Unpublished
2 (bases 1 to 1062)
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker
Cswy, Miami, FL 33129, USA
FEATURES
source
1..1062
/organism="Montastraea cavernosa"
/mol_type="mRNA"
/db_xref="taxon:63558"
1..82
83..766
/note="GFP-like; McsG1"
/codon_start=1
/product="green fluorescent protein G1"
/protein_id="AAU04446.1"
/db_xref="GI:51472051"
/translation="MSVTKPDMKIKLRMEGAVNGHKFVIEGDGKGPFEKGQIMDLTV
LEGPLPAPYDILTVFDYGNRVFAKYPKDI PDYFKOTFPEGYSWERSMTYEDQIC
ATNDITMKGVDVDCFEVKYIRFDGYNFPANGFVMDKTLKNEPSTEKMYRVDGLKGDV
NMALLEGGGHHYRCDFKTTYKARKVQLPDYFVDFHRIEIVSPDKYINKVLIYEHA
HFGLPROAK"
767..1062
3'UTR
polyA_signal 1016..1021
ORIGIN
Query Match 85.8%; Score 914.8; DB 2; Length 1062;
Best Local Similarity 92.6%; Pred. No. 8.7e-247;
Matches 989; Conservative 0; Mismatches 56; Indels 23; Gaps 2;
Qy      1  ATTCGCCCTGGTGTATTTGGAAGAGCGAGATCGAGAACACAAAGAGCTG--TAAGGTTGA 58
Db      1  ATTCACCTGTGTGATTTGGAAGAGAGCGAGATCGAGAACACAAAGAGCTGTTAAGGTTGA 60
Qy      59  TATCTTACTTACGCTTACCATCATGACAAAGTGTTCACAGGAAAGGGTGTGATTAAACC 118
Db      61  TATCTTACTTACGCTTACCCTCATGA-----GTGTGATAAAACC 99
Qy      119  AGACATGAAGATGAAGCTGCGTATGAAGAGGTGCTGTAAACGGGCACAAAGTTTCGTGTTGA 178
Db      100  AGACATGAAGATCAAGCTGCGTATGAAGAGGTGCTGTAAACGGGCACAAAGTTTCGTGTTGA 159

```

```

Qy      179  AGGAGATGGAAGGAGGAGCCCTTTTCGACGGAACACAGACTATGACCTTTACAGTCATAGA 238
Db      160  AGGAGACGGAAGGCAAGCCCTTTTCGAGGGGAAACAGACTATGACCTTTACAGTCATAGA 219
Qy      239  AGGCGCACCATTTGCGCTTTACGATATCTTGACAAAGTATTCGATTTACGGCAACAG 298
Db      220  AGGCGCACCATTTGCGCTTTACGATATCTTGACAAAGTATTCGATTTACGGCAACAG 279
Qy      299  GGTATTTCGCCCAATATCCCAAGAGACATAGCAGATTTTCAAGCAGACGTTTCTTGAGGG 358
Db      280  GGTATTTCGCCCAATATCCCAAGAGACATAGCAGATTTTCAAGCAGACGTTTTCGGAGGG 339
Qy      359  GTACTTCTCGGAAACGAAGCATGACATACGAGACCCAGGCGATTTGCATGCCCAACACGA 418
Db      340  GTACTTCTCGGAAACGAAGCATGACATACGAGACCCAGGCGATTTGCATGCCCAACACGA 399
Qy      419  CATACAAATGATGAAGGCGTTCGACGACTGTTTTCGCTATATAAATTCGATTTGATGTTGT 478
Db      400  CATACAAATGATGAAGGCGTTCGACGACTGTTTTCGCTATATAAATTCGATTTGATGTTGT 459
Qy      479  GAACTTTCTGCCCAATGTTCCAGTTATGCAGAGGAAGACGCTGAAATGGGAGCCATCCAC 538
Db      460  GAACTTTCTGCCCAATGTTCCAGTTATGCAGAGGAAGACGCTGAAATGGGAGCCATCCAC 519
Qy      539  TGAAGATAATGATGCGCGTGATGAGTGTCTGAAGCGTGATGTTAAACATGCTCTGTTGCT 598
Db      520  TGAAGATAATGATGCGCGTGATGAGTGTCTGAAGCGTGATGTTAAACATGCTCTGTTGCT 579
Qy      599  TGAAGAGGTGGCATACCGATGTCACTTCAAACTACTTACAAAGCTTAAGAGGTTGT 658
Db      580  TGAAGAGGTGGCATACCGATGTCACTTCAAACTACTTACAAAGCTTAAGAGGTTGT 639
Qy      659  CCGGTTTCCAGACTATACCTTTGTGACCATCGCTGAGATTTGTGAGATTTGTGAGCCACCAAGA 718
Db      640  CCGGTTTCCAGACTATACCTTTGTGACCATCGCTGAGATTTGTGAGATTTGTGAGCCACCAAGA 699
Qy      719  TTAACAACAGGTTAAGCTGCACGAGCATCGCAAGCTCGTCATGGAATGTCAAGGAAGGC 778
Db      700  TTAACAACAGGTTAAGCTGTATGAGCATCGCAAGCTCATTTTGGGCTGCGGAGGAGGC 759
Qy      779  CAAGTAAAGCTTAATGAAAAGTCAAGACGACAAACGAGGAGAACAAAGTACTTTTGT 838
Db      760  CAAGTAAAGCTTAATGAAAAGTCAAGACGACAAACGAGGAGAACAAAGTACTTTTGT 819
Qy      839  TAAATTTGAAGGCATTTACTCGGAATTTAGTATTGATTTGATCTTTTCGATTTCAAGGATTTGTC 898
Db      820  TAAATTTGAAGGCATTTACTCGGAATTTAGTATTGATTTGATCTTTTCGATTTCAAGGATTTGTC 879
Qy      899  CGGGATTTGTAGAGACTAGCTCTAGAGTTGTTATTTTGTGAAAAAAGATAGTTTCCAGTT 958
Db      880  CGAGATTTTANAGACTAGCTTANAGTTGTTATTTTGTGAAAAAANATAGTTTCCAGTT 939
Qy      959  TTTGCGGGATTAAGCATGGGATAGACTTTTAAACTCAGTTGTGTTGTTCAATTCAGTA 1018
Db      940  TTTGCGGGATTAAGCATGGGATAGACTTTTAAACTCAGTTGTGTTGTTCAATTCAGTA 999
Qy      1019  AGAAAACTGTAGTGAAGTAACTTGTATCGAAGCGCGAAAAA 1066
Db      1000  AGAAAACTGTAGTGAAGTAACTTGTATCGAAGCGCGAAAAA 1047

RESULT 5
AY181555
LOCUS Montastraea cavernosa isolate mc4 green fluorescent protein mRNA, complete cds.
DEFINITION
ACCESSION AY181555
VERSION AY181555.1 GI:32188171
KEYWORDS
SOURCE Montastraea cavernosa (great star coral)
ORGANISM Montastraea cavernosa
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Faviina; Faviidae; Montastraea.

```

REFERENCE 1 (bases 1 to 972)  
AUTHORS Kelmanson, I.V. and Matz, M.V.  
TITLE Molecular Basis and Evolutionary Origins of Color Diversity in Great Star Coral *Montastraea cavernosa* (Scleractinia: Faviida)  
JOURNAL Mol. Biol. Evol. 20 (7), 1125-1133 (2003)  
PUBMED 12777529

REFERENCE 2 (bases 1 to 972)  
AUTHORS Kelmanson, I.V. and Matz, M.V.  
TITLE Direct Submission  
JOURNAL Submitted (17-NOV-2002) Whitney Laboratory, University of Florida, 9505 Ocean Shore Blvd, St Augustine, FL 32080, USA

FEATURES  
Source Location/Qualifiers  
1..972  
/organism="Montastraea cavernosa"  
/mol\_type="mRNA"  
/isolate="mc4"  
/db\_xref="taxon:63558"  
77..760  
/codon\_start=1  
/product="green fluorescent protein"  
/protein\_id="AA061601.1"  
/db\_xref="GI:32188172"  
/translation="MSVIRPDMKIKLRMEGAVNGHKFVIEGDKGKPRGKOTMDLTV IEGLPFPAYDILTTVDYGNRVFAKPKDIPDYFKPTPEGSWERSMTYEDQGICI ATNDITMKGVDDCFYKIRFDGVNFPANGPKRTLEKWPSTERYVDRDGLKGDV NMALLIEGGGHYRCDKFTYTKAKVKVQLPDYHFDHRIEIVSHDKDYNKVKLYEHAE HSLPLPROAK"

CDS  
66 CTTTACCTCTACCATCATGACACAGTGTGACAGGAAAGGGTGTGATTAAACACAGACATG 125  
62 CTTTACCTCTACCATCATGACACAGTGTGACAGGAAAGGGTGTGATTAAACACAGACATG 100  
126 AGATGAAGCTGGTATGGAAGGTGCTGTAACGGGCAACAAGTTCGTGTTGAAAGGAGAT 185  
101 AAGATCAAGCTGCTATGGAAGGTGCTGTAACGGGCAACAAGTTCGTGTTGAAAGGAGAC 160  
186 GGAAGGAGGAGCCTTTCACGGNACACAGACTATGGACCTTACAGTCTATAGAGGCGCA 245  
161 GGAAGGAGGAGCCTTTCACGGGAAACACAGACTATGGACCTTACAGTCTATAGAGGCGCA 220  
246 CCATTGCGCTTTCGCTTACCATATCTTGACACAGTATTCGATTACGGCAACAGGGTATTC 305  
221 CTTTTCGCTTTCGCTTACCATATCTTGACACAGTATTCGATTACGGCAACAGGGTATTC 280  
306 GCCAATATCCCAAGACATACAGATATTTCAAGCAGACGCTTTCCTGAGGGGTACTTC 365  
281 GCCAATATCCCAAGACATACAGATATTTCAAGCAGACGCTTTCCTGAGGGGTACTTC 340  
366 TGGGACGAGGAGCATGACATACGAGNACGAGGCAATTTGCTATGCCACACACATACACA 425  
341 TGGGACGAGGAGCATGACATACGAGNACGAGGCAATTTGCTATGCCACACACATACACA 400  
426 ATGATGGAAGGCGTTCGACACTGTTTTCGCTATATAAATTCGATTGATGGTGTGAACCTT 485  
401 ATGATGGAAGGCGTTCGATGACTGTTTTCCTATATAAATTCGATTGATGGTGTGAACCTT 460  
486 CTGCAATATGCTCCAGTTATGACAGGAGAGACGCTGAATGGAGCCATCCACTGAGATA 545  
461 CTTGCCAATGCTCCAGTTATGACAGGAGAGACGCTGAATGGAGCCATCCACTGAGATA 520  
546 ATGATGCGCGTGTGAGTGTCTGAAGGGTGTATTTAATCGCTCTGTGCTTGAAGGA 605  
521 ATGATGCGCGTGTGAGTGTCTGAAGGGTGTATTTAATCGCTCTGTGCTTGAAGGA 580

ORIGIN  
Query Match 82.2%; Score 876; DB 2; Length 972;  
Best Local Similarity 93.9%; Pred. No. 7.8e-236;  
Matches 931; Conservative 0; Mismatches 40; Indels 21; Gaps 1;

QY 6 CCTGGTGTATTTGGAAGAGCAGATCGAGACAAACAAGAGCTGTGAAGTTCATCTTTA 65  
DB 2 CCTGGTGTATTTGGAAGAGCAGATCGAGACAAACAAGAGCTGTGAAGTTCATCTTTA 61

QY 66 CTTTACCTCTACCATCATGACACAGTGTGACAGGAAAGGGTGTGATTAAACACAGACATG 125  
DB 62 CTTTACCTCTACCATCATGACACAGTGTGACAGGAAAGGGTGTGATTAAACACAGACATG 100

QY 126 AGATGAAGCTGGTATGGAAGGTGCTGTAACGGGCAACAAGTTCGTGTTGAAAGGAGAT 185  
DB 101 AAGATCAAGCTGCTATGGAAGGTGCTGTAACGGGCAACAAGTTCGTGTTGAAAGGAGAC 160

QY 186 GGAAGGAGGAGCCTTTCACGGNACACAGACTATGGACCTTACAGTCTATAGAGGCGCA 245  
DB 161 GGAAGGAGGAGCCTTTCACGGGAAACACAGACTATGGACCTTACAGTCTATAGAGGCGCA 220

QY 246 CCATTGCGCTTTCGCTTACCATATCTTGACACAGTATTCGATTACGGCAACAGGGTATTC 305  
DB 221 CTTTTCGCTTTCGCTTACCATATCTTGACACAGTATTCGATTACGGCAACAGGGTATTC 280

QY 306 GCCAATATCCCAAGACATACAGATATTTCAAGCAGACGCTTTCCTGAGGGGTACTTC 365  
DB 281 GCCAATATCCCAAGACATACAGATATTTCAAGCAGACGCTTTCCTGAGGGGTACTTC 340

QY 366 TGGGACGAGGAGCATGACATACGAGNACGAGGCAATTTGCTATGCCACACACATACACA 425  
DB 341 TGGGACGAGGAGCATGACATACGAGNACGAGGCAATTTGCTATGCCACACACATACACA 400

QY 426 ATGATGGAAGGCGTTCGACACTGTTTTCGCTATATAAATTCGATTGATGGTGTGAACCTT 485  
DB 401 ATGATGGAAGGCGTTCGATGACTGTTTTCCTATATAAATTCGATTGATGGTGTGAACCTT 460

QY 486 CTGCAATATGCTCCAGTTATGACAGGAGAGACGCTGAATGGAGCCATCCACTGAGATA 545  
DB 461 CTTGCCAATGCTCCAGTTATGACAGGAGAGACGCTGAATGGAGCCATCCACTGAGATA 520

QY 546 ATGATGCGCGTGTGAGTGTCTGAAGGGTGTATTTAATCGCTCTGTGCTTGAAGGA 605  
DB 521 ATGATGCGCGTGTGAGTGTCTGAAGGGTGTATTTAATCGCTCTGTGCTTGAAGGA 580

QY 606 GTGGGCCATTACCGATGTGACTTCAAACTACTTACAAAGCTAAGAGGTTGTCCGGTTC 665  
DB 581 GTGGGCCATTACCGATGTGACTTCAAACTACTTACAAAGCTAAGAGGTTGTCCAGTTC 640

QY 666 CCAGACTATCATCTTTGTGGACCATCGCATTTGAGATTGTGAGCCACGACAAAGATTCAAC 725  
DB 641 CCAGACTATCATCTTTGTGGACCATCGCATTTGAGATTGTGAGCCACGACAAAGATTCAAC 700

QY 726 AAGGTTAAGCTGACGAGCATCGGAGCTCTCATGGACTGTCAAGGAGGCGCAAGTAA 785  
DB 701 AAGGTTAAGCTGATGAGCATCGGAGCTCTCATGGACTGTCAAGGAGGCGCAAGTAA 760

QY 786 AGGCTTAATGAAAAGTCAAGACGACAAACGAGCAGAAACAAAGTACTTTTTGTTAAATTT 845  
DB 761 AGGCTTAATGAAAAGTCAAGACGACAAACGAGCAGAAACAAAGTACTTTTTGTTAAATTT 820

QY 846 GAAGGCACTTACTCGGAATTAGTATTGATCTTTCGATTCGATTCGATTCGATTCGATTC 905  
DB 821 CAAGGCACTTACTCGGAATTAGTATTGATCTTTCGATTCGATTCGATTCGATTCGATTC 880

QY 906 TGTAGAGACTAGCTCTAGAGTGTGATTTTGTGAAGGAGGAGTTCAGTTCAGTTCGCGG 965  
DB 881 TGTAGAGACTAGCTCTAGAGTGTGATTTTGTGAAGGAGGAGTTCAGTTCAGTTCGCGG 940

QY 966 GATTACAGCATGGGATAGACTTTTTTAAACTC 997  
DB 941 GATTACAGCATGGGATAGACTTTTTTAAACTC 972

RESULT 6  
AY037768  
LOCUS AY037768 913 bp mRNA linear INV 31-MAY-2002  
DEFINITION Montastraea cavernosa clone 7.7 green fluorescent protein-like protein mRNA, complete cds.  
ACCESSION AY037768  
VERSION AY037768.1 GI:21303777  
KEYWORDS  
SOURCE Montastraea cavernosa (great star coral)  
ORGANISM Montastraea cavernosa  
Bukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia; Faviida; Faviidae; Montastraea.  
REFERENCE 1 (bases 1 to 913)  
AUTHORS Matz, M.V. and Lukyanov, S.A.  
TITLE Diversity and evolution of GFP-like fluorescent proteins  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 913)  
AUTHORS Matz, M.V. and Lukyanov, S.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Institute of Bioorganic Chemistry RAS, Miklukho-Maklaya 16/10, Moscow 117871, Russia  
FEATURES  
source Location/Qualifiers  
1..913  
/organism="Montastraea cavernosa"  
/mol\_type="mRNA"  
/db\_xref="taxon:63558"  
/clone="7.7"  
38..742  
/note="Fpncavgr7.7"  
/codon\_start=1  
/product="green fluorescent protein-like protein"  
/protein\_id="AAK71334.1"  
/db\_xref="GI:21303778"  
/translation="MTSVAQEKGVIKPDMKMLRMEGAVNGHKFVIEGDKGKPPDGT QTMDLTVEGLPFPAYLITTTVDYGNRVFAKPKDIPDYFKPTPEGSWERSMTY EDQIGCIATNDITMKGVDDCFYKIRFDGVNFPANGPKRTLEKWPSTERYVDRDGLKGDV NMALLIEGGGHYRDXKTTTAKVKVQLPDYHFDHRIEIVSHDKDYNKVK LYEHAEHSLPLPROAK"

ORIGIN  
Query Match 78.5%; Score 836.4; DB 2; Length 913;  
Best Local Similarity 96.2%; Pred. No. 1.2e-224;  
Matches 860; Conservative 6; Mismatches 27; Indels 1; Gaps 1;

52 AGGTTGATATCTTACTTACCTCTACCATCATGACAGTGTGTGCACAGAAAGGGTGTGA 111  
 9 AGGTTGATATCTTACTTACCTCTACCATCATGACAGTGTGTGCACAGAAAGGGTGTGA 68  
 112 TTAACACAGACATGAAGATGAAGCTGGTATGGAAGGTGCTGTAACCGGCACAAAGTTTC 171  
 69 TTAACACAGACATGAAGATGAAGCTGGTATGGAAGGTGCTGTAACCGGCACAAAGTTTC 128  
 172 TGGTTGAAGAGATGGAAGAGGAGGCTTTCGACGGAACACAGATATGACCTTACAG 231  
 129 TGATTGAAGAGATGGAAGAGGAGGCTTTCGACGGAACACAGATATGACCTTACAG 188  
 232 TCATAGAAGCGCACCATTCGCTTTCGCTTACGATATCTTGACAAAGTATTCGATTACG 291  
 189 TCATAGAAGCGCACCATTCGCTTTCGCTTACGATATCTTGACAAAGTATTCGATTACG 248  
 292 GCAACAGGGTATTGCGCAATATCCAGAAAGATGATAGCAGATATTTCAAGCAGACGTTTC 351  
 249 GCAACAGGGTATTGCGCAATATCCAGAAAGATGATAGCAGATATTTCAAGCAGACATTTTC 308  
 352 CTGAGGGGTACTTCTGGGAAGGAGGATGATAGCAGACGACGAGGCTTTCGATCGCCA 411  
 309 CTGAGGGGTACTTCTGGGAAGGAGGATGATAGCAGACGACGAGGCTTTCGATCGCCA 368  
 412 CAAACGACATAAATCATGATGAGGCGCTCGACGACTGTTTTCCTATAAAATTCGATTTC 471  
 369 CAAACGACATAAATCATGATGAGGCGCTCGACGACTGTTTTCCTATAAAATTCGATTTC 428  
 472 ATGGTGTGAATTTCTGCGCAATATGAGTATGAGAGGAGGAGCTGTAATGAGGAGC 531  
 429 ATGGTGTGAATTTCTGCGCAATATGAGTATGAGAGGAGGAGCTGTAATGAGGAGC 488  
 532 CATCCACTGAGATATGATGATGCGGCTGATGAGGAGGCTGTAATGAGGAGC 591  
 489 CATCCACTGAGAAATGATGATGCGGCTGATGAGGAGGCTGTAATGAGGAGC 548  
 592 TGTGCTTGAAGAGGAGGCTGATGAGGAGGCTGTAATGAGGAGC 651  
 549 TGTGCTTGAAGAGGAGGCTGATGAGGAGGCTGTAATGAGGAGC 608  
 652 AGGTTGTCAGGAGGCTGATGAGGAGGCTGTAATGAGGAGC 711  
 609 AGGTTGTCAGGAGGCTGATGAGGAGGCTGTAATGAGGAGC 668  
 712 ACAAGATATCAACAGGTTAAGCTGACGAGGAGGCTGTAATGAGGAGC 771  
 669 ACAAGATATCAACAGGTTAAGCTGATGAGGAGGCTGTAATGAGGAGC 728  
 772 GGAAGGCCAAGTAAAGGCTTAAAGGAGGCTGTAATGAGGAGC 830  
 729 GGAAGGCCAAGTAAAGGCTTAAAGGAGGCTGTAATGAGGAGC 788  
 831 TTTTGTGTTAAATTTGAAGGCTTAAAGGAGGCTGTAATGAGGAGC 890  
 789 TTTTGTGTTAAATTTGAAGGCTTAAAGGAGGCTGTAATGAGGAGC 848  
 891 ATTTGTTCCGGGATTTGTTAGAGGCTTAAAGGAGGCTGTAATGAGGAGC 944  
 849 ATTTGTTCCGGGATTTGTTAGAGGCTTAAAGGAGGCTGTAATGAGGAGC 902

RESULT 7  
 AF401282 1046 bp mRNA linear INV 05-AUG-2001  
 LOCUS Montastrea faveolata green fluorescent protein mRNA, complete cds.  
 DEFINITION AF401282  
 ACCESSION AF401282.1 GI:15081471  
 VERSION  
 KEYWORDS  
 SOURCE Montastrea faveolata  
 ORGANISM Montastrea faveolata  
 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
 Faviina; Faviidae; Montastrea.  
 REFERENCE 1 (bases 1 to 1046)

AUTHORS Lesser, M.P., Barry, T.M., Mazel, C., Matz, M.V., Lukyanov, S.A.,  
 Falkowski, P., Gorbunov, M. and Kolber, Z.  
 TITLE Green Fluorescent Proteins in Caribbean Scleractinian Corals  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1046)  
 AUTHORS Lesser, M.P. and Barry, T.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUL-2001) Department of Zoology and Center for Marine  
 Biology, University of New Hampshire, 46 College Road, Durham, NH  
 03824, USA

FEATURES  
 source  
 1. .1046  
 /organism="Montastrea faveolata"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:48498"  
 84. .767  
 /codon\_start=1  
 /product="green fluorescent protein"  
 /protein\_id="AAK83923.1"  
 /db\_xref="GI:15081472"  
 /translation="MSVLPKPMKIKLRMEGAVNGHKFVIEGDKGKPPFGTOSMDLTV  
 KEGALPFRAYDILITVDFYGNRVFAKYQDIPDPKQTFPEGSWERSMTYEDGICV  
 ATNDITLMKGVDDCFYKIRPDGVNFPANGPVNQKTKLKEPSTERYKVRDGLKGDV  
 NNALLIEGGGHYRCDFTTYAKKVFQVDPYHFVDRHIEILSHDKDYNKVLYEHAE  
 HSGLPQRAK"

ORIGIN  
 Query Match 78.3%; Score 834.2; DB 2; Length 1046;  
 Best Local Similarity 89.9%; Pred. No. 5e-224;  
 Matches 961; Conservative 0; Mismatches 78; Indels 30; Gaps 5;  
 QY 1 ATTGCGCCCTGGTGTGATTGGAAGAGGAGCAGATCGAGAAACAAAGAGCTG--TAAGGTGTA 58  
 Db 2 ATTGCGCCCTGGTGTGATTGGAAGAGGAGCAGATCGAGAAACAAAGAGCTGTTAAGGTTTA 61  
 QY 59 TATCTTACTTACCTTACCATCATGAAAGTGTGTCACAGGAAAGGGTGTGTTAAACC 118  
 Db 62 TATCTTACTTACCTTACCATCATGAAAGTGTGTCACAGGAAAGGGTGTGTTAAACC 100  
 QY 119 AGACATGAAGATGAAGCTCGTATGGAAGGCTGTAAAGCGGCACAAAGTTCGTGTTGA 178  
 Db 101 AGACATGAAGATGAAGCTCGTATGGAAGGCTGTAAAGCGGCACAAAGTTCGTGTTGA 160  
 QY 179 AGGAGATGAAAAGGGAAGCCTTTTCGACGGAACACAGACTATGAGCTTACAGTCATAGA 238  
 Db 161 AGGAGATGAAAAGGGAAGCCTTTTCGACGGAACACAGACTATGAGCTTACAGTCATAGA 220  
 QY 239 AGGCGCACCATTTGCCCTTTTCGATATCTTGAACACAGTATTCGATTCGGCAACAG 298  
 Db 221 AGGCGCGCTCTGCCCTTTTCGATATCTTGAACACAGTATTCGATTCGGCAACAG 280  
 QY 299 GGTATTCGCGCAATATCCGAGAGACATAGCAGATTTATTCAGCAGACGTTTCTCGAGG 358  
 Db 281 GGTATTCGCGCAATATCCGAGAGACATAGCAGATTTATTCAGCAGACGTTTCTCGAGG 340  
 QY 359 GTACTTCTGGGACGAGCATGACATACGAAAGACGAGGCAATTTGTCATGCCCAACACGA 418  
 Db 341 GTATTCCTGGGACGAGCATGACATACGAAAGACGAGGCAATTTGCGTCCGCAACACGA 400  
 QY 419 CATACCAATGATGGAAGGCGTCGACGCTGTTTTCGCTATATAAATTCGATTTGATGGTGT 478  
 Db 401 CATACCAATGATGGAAGGCGTCGACGCTGTTTTCGCTATATAAATTCGATTTGATGGTGT 460  
 QY 479 GAATTTCTTCCCAATGTCAGTATTCGAGAGGAGAGCTGAATGGAGGCATCCAC 538  
 Db 461 AATCTTCTTCCCAATGTCAGTATTCGAGAGGAGAGCTGAATGGAGGCATCCAC 520  
 QY 539 TGAGATAATGATGCGCGTGTGAGGTGCTGAGGAGTGTGTTAAATGCTCTGTGCT 598  
 Db 521 TGAGATAATGATGCGCGTGTGAGGTGCTGAGGAGTGTGTTAAATGCTCTGTGCT 580  
 QY 599 TGAAGGAGGTGGCCATTACCGATGTCATTCATAAACTACTTACAAAGCTTAAGAGGTGTT 658



REFERENCE 2 (bases 1 to 1045)  
 AUTHORS Kelmanson, I.V. and Matz, M.V.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-NOV-2002) Whitney Laboratory, University of Florida,  
 9505 Ocean Shore Blvd, St Augustine, FL 32080, USA

FEATURES  
 Location/Qualifiers  
 ..1045  
 /organism="Montastraea cavernosa"  
 /mol\_type="mRNA"  
 /isolate="mcl"  
 /db\_xref="taxon:63558"  
 85..762  
 /codon\_start=1  
 /product="red fluorescent protein"  
 /protein\_id="AA061598.1"  
 /db\_xref="GI:32188166"

CDS  
 85..762  
 /translation="MSVIVKSNVMIKLRMEGVSNGHNFVIVGEGKPKYEGTQSDMLTV  
 KEGAPLPFAFDITVTFVHGVNPFVAKYPKHIDPYFKMPPEGYSWERNFEDGGICT  
 ARNELTMEGDCRFNKFVDFVNFPPNGPVNOKTLKWPSTKMYVRDGLVLDINMA  
 LLEGGCHYRCDFRTTYRAKKGVKLPDYHFDVHSIELRHKDEYEVKLYEHAENHS  
 GLPRVAK"

ORIGIN  
 Query Match 63.8%; Score 679.6; DB 2; Length 1045;  
 Best Local Similarity 82.1%; Pred. No. 2.3e-180; Indels 44; Gaps 7;  
 Matches 883; Conservative 0; Mismatches 149;

QY 1 ATTGCGCCCTGGTGAATTGGAAGAGCAGATCGAGAACACAAAGAGCTG--TAAGGTTGA 58  
 DB 2 ATTCACCTGGTGAATTGGAAGAGCAGATCGAGAACACAAAGAGCTGATTAAGGCTGA 61  
 QY 59 TATCTTAC--TTAGCTTACCATCATGACAAAGTGTTCGACAGGAAGGGTGTGATTAAC 117  
 DB 62 TATCTTACTTTAGCTTACCATCATG-----GTGTGATTAAT 100  
 QY 118 CAGACATGAAGTGAAGTCTCGTATGGAAGTCTGTAAAGGGGCAAAAGTTCGTGGTTG 177  
 DB 101 CAGTATGAAGTCAAGTCTCGTATGGAAGGAGTGTAAAGGGGCAAAAGTTCGTGGTTG 160  
 QY 178 AAGGAGATGGAAGGAAGCCCTTTGAGGGAACACAGACTATGACCTTACAGTCAATAG 237  
 DB 161 TTGGAGAGGAGGAAGGCAAGCCCTTATGAGGGAACACAGAGTATGAGCTTACAGTCAAAG 220  
 QY 238 AAGGCGACCATTTGCCCTTACGATCTCTGACGATCTTGAACACAGTATTCGATTACGGCAAC 297  
 DB 221 AAGGCGACCTCTGCTCTGCTTACGATCTCTGACGATCTATGACACAGTATTCATACGGCAAT 280  
 QY 298 GGCTATTCGCAATACCCAGAGACATAGCAGATTTTCAAGCAGAGCTTTCTCTGAGG 357  
 DB 281 GGCTATTCGCAATACCCAGAGACATATCCAGACTATTCAGAGAGATGTTCTCTGAGG 340  
 QY 358 GGTACTTCTGGGAACGAGATGACATACGAGAGCAGGCAATTCATCGCCCAAAAG 417  
 DB 341 GGTATTCCTGGGAACGAGATGAAATTCGAGAGCGGGGCAATTCGACCGCAGGAAG 400  
 QY 418 ACATACATATGAGGAAGGCGTGAACGATGTTTGCCTATAAATTCGATTTGATGGTG 477  
 DB 401 AGATAACAATG-----GAAGGCGACTGTTTTCATTAAGTTCGATTTGATGGTG 451  
 QY 478 TGAATCTTCTGCAATGTTCCAGTTATGAGAGGAGAGCGCTGAATGGAGGCCATCCA 537  
 DB 452 TGAATCTTCTGCAATGTTCCAGTTATGAGAGGAGAGCGCTGAATGGAGGCCATCCA 511  
 QY 538 CTGAGATAATGATGCGCGTGTGAGTGTGAGGAGTGTGATGATGATGATGATGATGATG 597  
 DB 512 CTGAAAAATGATGTCGCGTGTGAGTGTGAGGAGTGTGATGATGATGATGATGATGATG 571  
 QY 598 TTGAAGAGGTGGCCATACCGATGACCTTCAAACTACTTCAAAAGCT---AAGAGG 654  
 DB 572 TTGAAGAGGTGGCCATACCGATGACCTTCAAACTACTTCAAAAGCT---AAGAGG 631  
 QY 655 TTGTCGGTTGCGAGACTATCACTTTCTGAGCCATCGCATTTGATGATTTGTCGAGCCAGCA 714

DB 632 GTGTCAAGTACCAGATTATCACTTTGTGGATCACTCCATTGAGATTTTGGCCATGACA 691  
 QY 715 AAGATTACAACAGGTTAAGCTGACGAGCATGCGAAGCTCGTCATCGACTGTCAAGA 774  
 DB 692 AAGAATACATGAGGTTAAGCTGTATGAGCATGCGAAGCTCATTTCTGGCTGCCAGGG 751  
 QY 775 AGGCCAAGTAAAGGCTTAATGAAAGTCAAGACGACAAAGAGGAGAAACAAAGTACTTTT 834  
 DB 752 TGGCAAGTAAAGGCTTAACGAAAGCCAGACCAAAAGAGAGAAAGAA--ATTTCG 809  
 QY 835 TTGTTAAATTTGAAGGCATTTACTCGGAATTAGTATTTGATCTTTCAATCAAGATTT 894  
 DB 810 TTTTCTATTGAAAGGCAATTTACTCGGAATTAGTATTTGATCTTTCAATCAAGGTTT 869  
 QY 895 GTTCGGGATTTGTAGAGACTAGCTCTA-----GAGTTGATTTTGTGAAAGAGATA 948  
 DB 870 GTTACGGACTTTGTGGAGACTAGCTCTTAAACAACATGACTAATGAAGTTTGTGGAAGAAAG 929  
 QY 949 GTTTCAGTGTTCGGGATTTACAGCATGGGATAGACTTTTAAACACTCAGTTGTGGTCA 1008  
 DB 930 TTTCCAAATTTTGTGATTTACAGCATAGATAGATTTCTTTAACTCAGTTGTGGAA 989  
 QY 1009 AATGCAAGTAAAGAACTGTAGTGAATAAACTTTGATCGAAGCCGAAAGAAAA 1064  
 DB 990 AATCAAGTAAAGAAATATAGTGAATAAAATTTGTTGCGAAGCCAGAAAAAAA 1045

RESULT 10  
 AY181553 836 bp mRNA linear INV 25-JUN-2003  
 LOCUS Montastraea cavernosa isolate mc2 green fluorescent protein mRNA,  
 complete cds.  
 ACCESSION AY181553  
 VERSION AY181553.1 GI:32188167  
 KEYWORDS  
 SOURCE Montastraea cavernosa (great star coral)  
 ORGANISM Montastraea cavernosa  
 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
 Faviina; Faviidae; Montastraea.  
 REFERENCE 1 (bases 1 to 836)  
 Kelmanson, I.V. and Matz, M.V.  
 Molecular Basis and Evolutionary Origins of Color Diversity in  
 Great Star Coral Montastraea cavernosa (Scleractinia: Faviida)  
 JOURNAL Mol. Biol. Evol. 20 (7), 1125-1133 (2003)  
 PUBMED 12777529  
 REFERENCE 2 (bases 1 to 836)  
 Kelmanson, I.V. and Matz, M.V.  
 Direct Submission  
 TITLE Submitted (17-NOV-2002) Whitney Laboratory, University of Florida,  
 9505 Ocean Shore Blvd, St Augustine, FL 32080, USA  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1..836  
 /organism="Montastraea cavernosa"  
 /mol\_type="mRNA"  
 /isolate="mc2"  
 /db\_xref="taxon:63558"  
 84..767  
 /codon\_start=1  
 /product="green fluorescent protein"  
 /protein\_id="AA061599.1"  
 /db\_xref="GI:32188168"  
 /translation="MSVIVKSNVMIKLRMEGVSNGHNFVIVGEGKPKYEGTQSDMLTV  
 KEGAPLPFAFDITVTFVHGVNPFVAKYPKHIDPYFKMPPEGYSWERNFEDGGICTV  
 ATNDITLMKGDDCFVYKIRFDGVNFPANGPVNOKTLKWPSTKMYVRDGLVLDGVDV  
 NNALLGEGGHYRCDFRTTYRAKKGVKLPDYHFDVHSIELRHKDEYEVKLYEHAENHS  
 HSGLPRAK"

CDS  
 84..767  
 /codon\_start=1  
 /product="green fluorescent protein"  
 /protein\_id="AA061599.1"  
 /db\_xref="GI:32188168"

ORIGIN  
 Query Match 63.6%; Score 678.4; DB 2; Length 836;  
 Best Local Similarity 89.5%; Pred. No. 4.9e-180; Indels 23; Gaps 2;  
 Matches 761; Conservative 0; Mismatches 66;

QY 1 ATTGCGCCCTGGTGAATTGGAAGAGCAGATCGAGAACACAAAGAGCTG--TAAGGTTGA 58





QY 659 CCGTGTCCAGACTATCACTTTGTGGACCATCGCAATTCAGATTGTGAGCCACGACAGA 718  
 Db 729 CCAGATCCAGACTATCACTTTGTGAATCAACCGACTTGAGATAACATGGCATGACGAGGA 788  
 QY 719 TTACAAACAAGTTAAGCTGACGAGCATGCCGAAGCTCGTCATGAGCTGTCAAGGAAGGC 778  
 Db 789 TTACAAACAAGTTAAGCTGCTGAGCATGCGAAGCTCATCTGAGCTGCCAGGACGC 848  
 QY 779 CAAGTAAGGCTTAATGAAAGTCAAGACGCAACAGGAGGAACA -----AAGTA 829  
 Db 849 CAATAAAGGCTTGACGAAAGCCAAACCGCAAGAGTCAAGAAAGTATATATAATG 908  
 QY 830 CTTTTTTTAAATTTGAAGCAATTTACTCGAATTTAGTATTTCATCTATCTTCGATTCAG 889  
 Db 909 TATATTTTTCAACTGAAGGCAATTTCCACTCGGAATTTAGTATTTCATCTTCGATTCAG 968  
 QY 890 GATTGTTCGGGATTTGTAGACACTAGCTCTAGAGTTGTATTTTGTGAAAAAGATAG 949  
 Db 969 GATTATTTGGGATTTGCTAGCCACTAGCTTTATTTGTTAAATTAAGTTAAAGAGC---G 1025  
 QY 950 TTTCCAGTTTTCGGGATTTACAGATGCGGATAGCTTTTAAACTCAGTTGTGTCAA 1009  
 Db 1026 TTTAGCATTTTTCGGTATTACAACATAGGCACAGAGCTCTTAACCCAGTAGTGTGTGAG 1085  
 QY 1010 ATGCAAGTAAAGAACTGTAGTGAATTAACCTTTGTTTCGAA 1052  
 Db 1086 GTACAAGTAAAGAACTTTGTTGAGATAGACTTTGTAGTCGAA 1128

RESULT 12  
 AF384683 1046 bp mRNA linear INV 27-AUG-2001  
 LOCUS Montastrea cavernosa green fluorescent protein mRNA, complete cds.  
 DEFINITION  
 ACCESSION AF384683  
 VERSION AF384683.2 GI:15298095  
 KEYWORDS  
 SOURCE Montastrea cavernosa (great star coral)  
 ORGANISM Montastrea cavernosa  
 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
 Faviina; Faviidae; Montastrea.  
 1 (bases 1 to 1046)  
 Lesser, M.P., Barry, T.M., Mazel, C., Matz, M.V., Lukyanov, S.A.,  
 Falkowski, P., Gorbunov, M. and Kolber, Z.  
 Green fluorescent proteins in Caribbean Scleractinian corals  
 Unpublished  
 2 (bases 1 to 1046)  
 Lesser, M.P. and Barry, T.M.  
 Direct Submission  
 Submitted (21-MAY-2001) Department of Zoology and Center for Marine  
 Biology, University of New Hampshire, 46 College Road, Durham, NH  
 03824, USA  
 3 (bases 1 to 1046)  
 Lesser, M.P. and Barry, T.M.  
 Direct Submission  
 Submitted (27-AUG-2001) Department of Zoology and Center for Marine  
 Biology, University of New Hampshire, 46 College Road, Durham, NH  
 03824, USA  
 REMARK  
 COMMENT Sequence update by submitter  
 On Aug 27, 2001 this sequence version replaced gi:14485239.  
 FEATURES  
 source  
 1..1046  
 /organism="Montastrea cavernosa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:63558"  
 89..766  
 /note="GFP"  
 /codon\_start=1  
 /product="green fluorescent protein"  
 /protein\_id="AAK62982.2"  
 /db\_xref="GI:15298096"  
 /translation="MSVTKPIEIKRMQGVVNGHKFVKGEGKPEFGQTINLTV  
 KEGAPLPAYDILISAFOYGNRVFTKYEDDIPDYFKQTFPPGYSWERIMAYEDQSICT

ATSDIKMEGDGCFIYIEIOFHGVNFPNPGVMQKTKLWEPSTKMYVRDGVLYKDVNNA  
 LILLEGSHYRCDPRSTYKAKKRVQLPDYHFVDRIBILSHDNDYNTVTKLSEDAEAYS  
 MLPSQAK"

Query Match 58.7%; Score 625.4; DB 2; Length 1046;  
 Best Local Similarity 80.2%; Pred. No. 4.7e-165;  
 Matches 875; Conservative 0; Mismatches 141; Indels 75; Gaps 9;

QY 1 ATTGCGCCCTGGTGAATTTGGAAGAGCAGATCGAAGCAACAAGAGCTG--TAAGGTGA 58  
 Db 2 ATTGCGCCCTGGTGAATTTGGAAGAGCAGATCGAAGCAACAAGAGCTGTATAGGTGA 61  
 QY 59 TATCTTACT-----TACGTCTACCATCATGCAAGTGTTCACAGGAAAAAGGCTGTGATT 113  
 Db 62 TACTGTATCTTAGTATAGCTTACCATCTGA-----GTGTGATT 100  
 QY 114 AAAACGAGCATGAAGATGAAGCTGCGTATGGAAGTGTGTAACCGGGCAACAAGTTCGTG 173  
 Db 101 AAACCAATCATGAGATCAAGCTGCGTATGCAAGGCGTTGTAACCGGGCAACAAGTTCGTG 160  
 QY 174 GTTGAAGGAGATGGAAGAGGAGCCCTTCGACGGAACACAGACTATGGAACCTTACAGTC 233  
 Db 161 ATTAAGGAGAGGAGGAGGAGGCAAGCCCTTCGAGGGAACGAGACTATAAACCTTACAGTC 220  
 QY 234 ATAGAAGGCGCACCATTTGCTTTCGCTTACGATATCTTGACAAACAGTATTCGATTACGGC 293  
 Db 221 AAAGAAGGCGCACCTCTCCCTTTTGTTCAGCATCTTGACATCAGATTCCAGTATGGT 280  
 QY 294 AACAGGTTATTCGCAAAATACCCAGAAAGACATAGCAGATTATTTCAAGCAGAGAGCTTTCCT 353  
 Db 281 AACAGGTTATTCACCAATATATCCAGACGATATACAGACTATTTTCAAGCAGAGAGCTTTCCT 340  
 QY 354 GAGGGTACTTTCGGAAACGAGCATGACATACGAGACCGAGGCAATTCGATCGCCACA 413  
 Db 341 GAAGGGTATTTCGTGGGAGCGAATCATGGCTTATGAAGACCCAGAGATTTTGCACGGCCACA 400  
 QY 414 AACGACATAACAATGATGGAAGGCGTCGACGACTGTTTTTGCCTATATAAATTCGATTTCAT 473  
 Db 401 AGCAGCATATAAATG-----GAAGGCGACTGTTTATCTACGAAATTCATTTTCAT 451  
 QY 474 GGTGTGAACCTTTCCTGCCAATGTCAGTTATCGAGGAAGACGCTGAAATGGGAGCCA 533  
 Db 452 GGTGTGAACCTTTCACCCCAATGTCAGTTATCGAGGAAGACGCTGAAATGGGAGCCA 511  
 QY 534 TCACCTGAGTATATGTCGCGTGTGAGTGTGAGGCTGATGTTACATGGGCTCTG 593  
 Db 512 TCCACTGAGAAATGATGTGCGTGTGAGTGTGAGGCTGATGTTACATGGCTCTG 571  
 QY 594 TTGCTTGAAGGAGGTGGCCATTACCGATGTGAATTCGAAACTACTTTACAAAGCTTAAGAAG 653  
 Db 572 TTGCTTGAAGGAGGTGGCCATTACCGATGTGAATTCGAAAGTACTTTACAAAGCGAGAGAA 631  
 QY 654 GTTGTGCGGTTGCCAGATATCACTTTTGGACCATCGCATTCGAGATTGTGAGCCAGAC 713  
 Db 632 CGTGTTCAGTTGCCAGACTATCACTTTTGTGGACCCAGCATTTGAGATTTTGGAGCCATGAC 691  
 QY 714 AAAGTTTACAAAGTTTGAAGTTCGACGACATGCGAAGCTCG---TCATGGACTCTGCA 770  
 Db 692 AATGACTACAAACCGGTAAGCTGTCTGAGGATGCCGAGGCTCGCTATTTCTATCTGCTGCCG 751  
 QY 771 AGGAAGGCCCAAGTAAAGGCTTAATGAAAGTCAAGACGACAAACGAGGAGAGAAACAAA---- 826  
 Db 752 AGTCAGGCCCAAGTAAAGGCTTAACGAAAGGCTTAACGACAAACGAGGAGAGAAAGAGTG 811  
 QY 827 -----GTACTTTTTTGTAAATTTGAAGGCAATTTACTTCGGAATTTAGTATT 871  
 Db 812 TTTTGTAGTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 871  
 QY 872 TGATACTTTTCGATTCAAGGATTTTGTCCGGGATTTGTTAGAGACTAGCTCTAGAGTTGTA 931  
 Db 872 GAATA-TTTCGATCCAGGGATTT-----GCGATTTGTTAGACACTAGCTCTAGAGTTAAA 925



```

QY 932 TTTTGTGAAAAAGATAGTTCCTTTCAGTTTTCCTCGGGATTACAGCATGGGATGACTTTTT 991
DB 926 GTTTATGAAAAAAGTTT-----CGGGATTCCATCACAAGAATAGACTTTTT 975

QY 992 AAACCTCAGTTGGTCAATGCAAGTAAAGAAACTAGTAGAATAAACTTTGTTATCGA 1051
DB 976 AAACCTCAGTTGTAGACAAAAACAAGTAAGAAACTTTAAGTGAATAAACTTTGTTATCGA 1035

QY 1052 AGCCGAAAAAA 1062
DB 1036 AGCCGAAAAAA 1046

RESULT 13
LOCUS AY679112 684 bp mRNA linear INV 25-AUG-2004
DEFINITION Montastraea faveolata green fluorescent protein G1 mRNA, complete
cds.
ACCESSION AY679112
VERSION AY679112.1 GI:51472056
KEYWORDS Montastraea faveolata
SOURCE Montastraea faveolata
ORGANISM Montastraea faveolata
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
Favilina; Faviidae; Montastraea.
REFERENCE 1 (bases 1 to 684)
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.
TITLE Cloning of Chnidarian Fluorescent Protein Genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 684)
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker
Cswy, Miami, FL 33129, USA
FEATURES
source
location/Qualifiers
1..684
/organism="Montastraea faveolata"
/mol_type="mRNA"
/db_xref="taxon:48498"
1..684
/note="GFP-like; MfaG1"
/codon_start=1
/product="green fluorescent protein G1"
/protein_id="AAU04449.1"
/db_xref="GI:51472057"
/translation="MSVVKPDMKIKLRMEGAVNGHKFVIEGDGKGPPEGKQWDLTV
IEGAPLPAYDILTTVDYGNRVFAKYPKIDIPYKQTFPEGSWERSMTYEDQIC
ATNDITMKGVDDCFVKIRDFGNFPANGPVMQRTKLKWFSTKMYVRDGLKGDV
NMALLGGGHHYRCDSTTKYKAKVQVLPDYHFDVHRIEIVSHDKDYNKVKLYEHAE
HSLGLPROAK"
ORIGIN
Query Match 58.6%; Score 624.8; DB 2; Length 684;
Best Local Similarity 94.6%; Pred. No. 6.6e-165;
Matches 647; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 102 AAGGTGTGATTAAACACAGATGAGATGAGCTGCGTAGGAGGTGCTGTAAACGGG 161
DB 1 ATGAGTGTGATAAAACACAGATGAGATGAGCTGCGTAGGAGGTGCTGTAAACGGG 60

QY 162 CACAAAGTTCGTGGTTGAAGGAGATGGAAGGAAGCGCTTTCGACGGAACACAGACTATG 221
DB 61 CACAAAGTTCGTGGTTGAAGGAGATGGAAGGAAGCGCTTTCGACGGAACACAGACTATG 120

QY 222 GACCTTACAGTCATAGAGGGCGACCATTCGCTTTCGCTTACGATATCTTGACACAGTA 281
DB 121 GACCTTACAGTCATAGAGGGCGACCATTCGCTTTCGCTTACGATATCTTGACACAGTT 180

QY 282 TTCGATTACGCAACAGAGGTATTCGCCAAATACCCAGAGACATAGAGATTATTTCAAG 341
DB 181 TTCGATTACGCAACAGAGGTATTCGCCAAATACCCAGAGACATAGAGATTATTTCAAG 240

QY 342 CAGACGTTTCCTGAGGGGTACTTCTGGGAAACGAAGCATGACATACGAAGACAGGGCATT 401

```

```

DB 241 CAGACGTTTCCTGAGGGGTACTCTCTGGGAAACGAAGCATGACATACGAAGACAGGGCATT 300
QY 402 TGCATCGCCACAAACGACATAACATGATGAGAGCGCTCGACGACTGTTTTCCTATAAA 461
DB 301 TGCATCGCCACAAACGACATAACATGATGAGAGCGCTCGACGACTGTTTTCCTATAAA 360
QY 462 ATTCGATTTCGATGGTGTGAACTTTCTCCTCAATGTCAGTTTATCGAGGAGAGCGGTG 521
DB 361 ATTCGATTTCGATGGTGTGAACTTTCTCCTCAATGTCAGTTTATCGAGGAGAGCGCTA 420
QY 522 AAATGGGAGCCATCCACTGAGATAATGATATGCGCGTGTAGAGTCTGAAGGTTGATGTT 581
DB 421 AAATGGGAGCCATCCACTGAGAAAAAATGATGTCGCGTGTAGGCTACTGAAGGTTGATGTT 480
QY 582 AACATGGCTCTGTTGCTTGAAGGAGGTGGCCATTACCGATGACTTCAAACTACTTAC 641
DB 481 AACATGGCTCTGTTGCTTGAAGGAGGTGGCCATTACCGATGACTTCAAACTACTTAC 540
QY 642 AAAGCTAAGAAAGTTGTCGGTTCGACAGACTATCACTTTGTGGACATCGCATTCGAGATT 701
DB 541 AAAGCTAAGAAAGTTGTCGGTTCGACAGACTATCACTTTGTGGACATCGCATTCGAGATT 600
QY 702 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGACGAGCATGCCGAAAGCTCGTCAT 761
DB 601 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGATGAGCATGCCGAAAGCTCATCTT 660
QY 762 GCAGCTGTCAAGGAAGGCCAAGTAA 785
DB 661 GGGCTGCCGAGGAGGCCCAAGTAA 684

RESULT 14
LOCUS AY679110 684 bp mRNA linear INV 25-AUG-2004
DEFINITION Montastraea cavernosa enhanced green fluorescent protein Glea mRNA,
complete cds.
ACCESSION AY679110
VERSION AY679110.1 GI:51472052
KEYWORDS Montastraea cavernosa (great star coral)
SOURCE Montastraea cavernosa
ORGANISM Montastraea cavernosa
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
Favilina; Faviidae; Montastraea.
REFERENCE 1 (bases 1 to 684)
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.
TITLE Cloning of Chnidarian Fluorescent Protein Genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 684)
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker
Cswy, Miami, FL 33129, USA
FEATURES
source
location/Qualifiers
1..684
/organism="Montastraea cavernosa"
/mol_type="mRNA"
/db_xref="taxon:63558"
1..684
/note="GFP-like; McaGlea"
/codon_start=1
/product="enhanced green fluorescent protein Glea"
/protein_id="AAU04447.1"
/db_xref="GI:51472053"
/translation="MSVVKPDMKIKLRMEGAVNGHKFVIEGDGKGPPEGKQWDLTV
IEGAPLPAYDILTTVDYGNRVFAKYPKIDIPYKQTFPEGSWERSMTYEDQIC
ATNDITMKGVDDCFVKIRDFGNFPANGPVMQRTKLKWFSTKMYVRDGLKGDV
NMALLGGGHHYRCDSTTKYKAKVQVLPDYHFDVHRIEIGSPDKDYNKVKLYEHAE
HFGLPROAK"
ORIGIN
Query Match 58.5%; Score 623.2; DB 2; Length 684;
Best Local Similarity 94.4%; Pred. No. 1.9e-164;

```



Search completed: April 5, 2006, 21:46:56  
Job time : 3834 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 20:42:49 ; Search time 766 Seconds  
(without alignments)  
9274.886 Million cell updates/sec

Title: US-10-757-356-17  
Perfect score: 1066  
Sequence: 1 attcgccctgggtatttga.....atcggaagccgaaaaaaa 1066

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_21.\*

- 1: geneseqn1980s.\*
- 2: geneseqn1990s.\*
- 3: geneseqn2000s.\*
- 4: geneseqn2001as.\*
- 5: geneseqn2001bs.\*
- 6: geneseqn2002as.\*
- 7: geneseqn2002bs.\*
- 8: geneseqn2003as.\*
- 9: geneseqn2003bs.\*
- 10: geneseqn2003cs.\*
- 11: geneseqn2003ds.\*
- 12: geneseqn2004as.\*
- 13: geneseqn2004bs.\*
- 14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1066	100.0	1066	8	AAL60081 Montastra
2	838.8	78.7	913	8	AAL60085 Montastra
3	665	62.4	1133	8	AAL60086 M. annula
4	628	58.9	684	10	ADH00980 Montastra
5	628	58.9	684	12	ADO78073 M. cavern
6	628	58.9	684	14	ADZ84228 Montastra
7	626.4	58.8	684	12	ADO78071 M. cavern
8	626.4	58.8	684	14	ADZ84226 Montastra
9	540	50.7	684	14	ADW43033 Montastra
10	540	50.7	684	14	ADW43022 Favia fav
11	536.8	50.4	684	14	ADW43035 Favia fav
12	533.6	50.1	684	14	ADW43041 Favia fav
13	533.4	50.0	795	8	AAL60080 Montastra
14	527.2	49.5	684	14	ADW43037 Favia fav
15	527.2	49.5	684	14	ADW43039 Favia fav
16	517.2	48.5	860	10	ABZ76493 M. cavern
17	507.4	47.6	729	14	ADY53076 Great sta
18	507.2	47.6	678	14	ADY53068 Great sta
19	499.2	46.8	681	13	ADU21136 M. cavern

20	497.8	46.7	729	14	ADY53077 Great sta
21	497.4	46.7	678	14	ADY53069 Great sta
22	486.4	45.6	678	10	ADD67209 Galaxea f
23	467.2	43.8	675	14	ADY53067 Great sta
24	467.2	43.8	726	14	ADY53075 Great sta
25	465	43.6	726	10	ADH00962 Montastra
26	465	43.6	1111	6	ABK96463 cDNA frag
27	462	43.3	1113	6	ABK96462 cDNA frag
28	461.8	43.3	726	10	ADH00964 Montastra
29	460.2	43.2	726	10	ADH00972 Montastra
30	458.6	43.0	726	10	ADH00970 Montastra
31	457	42.9	726	10	ADH00968 Montastra
32	455.4	42.7	726	10	ADH00966 Montastra
33	436	40.9	745	10	ADH00976 Montastra
34	436	40.9	746	10	ADH00974 Montastra
35	435.2	40.8	675	14	ADW21090 Montastra
36	433	40.6	690	12	ADL72772 Lophophytu
37	433	40.6	748	10	ADH00978 Montastra
38	432.8	40.6	684	12	ADN62702 Reporter
39	431.6	40.5	681	10	ADH00960 Montastra
40	430.2	40.4	678	12	ADL72766 T. geoffr
41	429.8	40.3	690	12	ADL72774 S. vittien
42	429.8	40.3	690	12	ADL72770 S. vittien
43	425.4	39.9	678	12	ADL72768 T. geoffr
44	421.2	39.5	1116	3	AAAS2766 Clavulari
45	421.2	39.5	1116	4	AAO03611 Clavulari

ALIGNMENTS

RESULT 1  
AAL60081  
ID AAL60081 standard; DNA; 1066 BP.  
XX  
AC AAL60081;  
XX  
DT 27-AUG-2003 (first entry)  
XX

DE Montastra cavernosa green fluorescent protein (mcavGFP) encoding DNA.  
XX Chromoprotein; fluorescent protein; colouring agent; food composition;  
XX cosmetic; fluorescence resonance energy transfer; calcium ion indicator;  
KW biosensor; pH indicator; phosphorylation indicator; messenger detector;  
KW transgenic animal; fluorescent timer; green fluorescent protein; GFP;  
KW transgenic; gene; ds.  
XX Montastra cavernosa.  
XX

Key Location/Qualifiers  
CDS 81..785  
FT /\*tag= a  
FT /product= "mcavGFP protein"  
XX

WO2003042401-A2.

22-MAY-2003.

12-NOV-2002; 2002WO-US036499.

13-NOV-2001; 2001US-0332980P.

(CLON-) CLONTECH LAB INC.

Labas YA, Gurskaya NG, Yanushevich Y, Fradkov AF, Lukyanov K;

Lukyanov S, Matz MV;

WFI; 2003-482353/45.

P-PSDB; AAO29586.

New chromo and/or fluorescent-protein encoding nucleic acids, the proteins being useful as labels in analyte detection assays, as selectable markers in recombinant DNA application, or as biosensors in

prokaryotic and eukaryotic cells.

Claim 1; Fig 16; 89pp; English.

The invention relates to novel chromo and/or fluorescent proteins and nucleic acid molecules encoding such proteins. Sequences of the invention are useful as labelling tools for marking a protein, cell or organism in biochemistry, molecular biology and medical diagnostic applications. Chromoproteins are useful as colouring agents capable of imparting colour or pigment to a particular composition of matter, e.g. food compositions, pharmaceuticals or cosmetics, as labels in analyte detection assays, or as selectable markers in recombinant DNA applications. Fluorescent proteins are useful in fluorescence resonance energy transfer applications, in high-throughput screening assays, as biosensors in prokaryotic and eukaryotic cells e.g. as calcium ion indicator, as pH indicator or as phosphorylation indicator, as second messenger detectors, as in vivo markers in transgenic animals, or as fluorescent timers. The present sequence is Montastraea cavernosa green fluorescent protein (mcaVgFP) encoding DNA

Sequence 1066 BP; 336 A; 188 C; 272 G; 270 T; 0 U; 0 Other;

Query Match 100.0%; Score 1066; DB 8; Length 1066;

Best Local Similarity 100.0%; Pred. No. 5.6e-313;

Matches 1066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATTGCGCCCTGGTGAATTTGGAGAGAGAGAGATCGAGAACCAACAGAGCTGTAAAGTTGATA	60
Db	1	ATTGCGCCCTGGTGAATTTGGAGAGAGAGAGATCGAGAACCAACAGAGCTGTAAAGTTGATA	60
Qy	61	TCCTTACTTACGCTACCATCATGACAGTGTGACAGGAAAGGGTGTGATTAACACGAG	120
Db	61	TCCTTACTTACGCTACCATCATGACAGTGTGACAGGAAAGGGTGTGATTAACACGAG	120
Qy	121	ACATGAAGATGAAGCTGGTATGGAAGGTGCTGTAACGGGCACAAAGTTCGTGTTCAAG	180
Db	121	ACATGAAGATGAAGCTGGTATGGAAGGTGCTGTAACGGGCACAAAGTTCGTGTTCAAG	180
Qy	181	GAGATGGAAGGAGGAGCTTTCACGGAACACAGATGTCGACTTACAGTCTATAGAG	240
Db	181	GAGATGGAAGGAGGAGCTTTCACGGAACACAGATGTCGACTTACAGTCTATAGAG	240
Qy	241	CGCGACCATTCCTTTCCTTACGATATCTTGAACAGATTCGATTCGCGCAACAGGG	300
Db	241	CGCGACCATTCCTTTCCTTACGATATCTTGAACAGATTCGATTCGCGCAACAGGG	300
Qy	301	TATTGCGCAATACCCAGAGACATAGCAGATTTATTCAGCAGACGTTTCCCTGAGGGT	360
Db	301	TATTGCGCAATACCCAGAGACATAGCAGATTTATTCAGCAGACGTTTCCCTGAGGGT	360
Qy	361	ACTTCTGGGAACGAAGCATGATACGAGACCGAGGCGATTTGCATCGCCACCAACGACA	420
Db	361	ACTTCTGGGAACGAAGCATGATACGAGACCGAGGCGATTTGCATCGCCACCAACGACA	420
Qy	421	TAAACAATGATGGAAGGCTGCACGACTGTTTTCCTATATAAATTCGATTCGTTGTA	480
Db	421	TAAACAATGATGGAAGGCTGCACGACTGTTTTCCTATATAAATTCGATTCGTTGTA	480
Qy	481	ACTTTCCTGCCAATGGTCCAGTTATGACAGAGAAAGACGCTGAATGGAGGCCATCCACTG	540
Db	481	ACTTTCCTGCCAATGGTCCAGTTATGACAGAGAAAGACGCTGAATGGAGGCCATCCACTG	540
Qy	541	AGATAATGATGCGGTGATGAGTGTGATGAGGTTGATTTAACTGCTCTGTTGTTG	600
Db	541	AGATAATGATGCGGTGATGAGTGTGATGAGGTTGATTTAACTGCTCTGTTGTTG	600
Qy	601	AAGAGGTGGCCATTACCGATGCTCTTCAAACTTACATCAAGAGCTAAGAGGTTGTC	660
Db	601	AAGAGGTGGCCATTACCGATGCTCTTCAAACTTACATCAAGAGCTAAGAGGTTGTC	660
Qy	661	GTTTGCCAGACTATCACCTTTGTGGACCATCGCATTTGAGATTTGAGCCACCAAGATT	720
Db	661	GTTTGCCAGACTATCACCTTTGTGGACCATCGCATTTGAGATTTGAGCCACCAAGATT	720

Qy	721	ACAACAAGGTTAAGCTGCAGCATGCCGAGCTCGTATGACTGTCAAGGAAGGCCA	780
Db	721	ACAACAAGGTTAAGCTGCAGCATGCCGAGCTCGTATGACTGTCAAGGAAGGCCA	780
Qy	781	AGTAAAGGCTTAATGAAAAAGTCAAGACGACCAACGAGGAGAAACAAAGTACTCTTTT	840
Db	781	AGTAAAGGCTTAATGAAAAAGTCAAGACGACCAACGAGGAGAAACAAAGTACTCTTTT	840
Qy	841	AATTGAAGCATTTACTCGGAATTAGTATTGATTAATTCGATTCAAGGATTGTTCCG	900
Db	841	AATTGAAGCATTTACTCGGAATTAGTATTGATTAATTCGATTCAAGGATTGTTCCG	900
Qy	901	GGATTTGTTAGACATGCTAGAGTGTGATTTCTGAAAAAGAGATGTTCCAGTTT	960
Db	901	GGATTTGTTAGACATGCTAGAGTGTGATTTCTGAAAAAGAGATGTTCCAGTTT	960
Qy	961	TGCGGATTTACAGCATGGGGATAGACTTTTAACTCAGTTGTTGGTCAAAATGCAAGTAAG	1020
Db	961	TGCGGATTTACAGCATGGGGATAGACTTTTAACTCAGTTGTTGGTCAAAATGCAAGTAAG	1020
Qy	1021	AAAACCTGATGAGATAAACTTGTTCGAAAGCCGAAAAA	1066
Db	1021	AAAACCTGATGAGATAAACTTGTTCGAAAGCCGAAAAA	1066

RESULT 2

AAAL60085  
ID AAL60085 standard; DNA; 913 BP.

XX AC AAL60085;

DT 27-AUG-2003 (first entry)

XX Montastraea cavernosa green fluorescent protein (mcaVgFP) 2 encoding DNA.

XX Chromoprotein; fluorescent protein; colouring agent; food composition;  
cosmetic; fluorescence resonance energy transfer; calcium ion indicator;  
biosensor; pH indicator; phosphorylation indicator; messenger detector;  
transgenic animal; fluorescent timer; green fluorescent protein; GFP;  
transgenic; gene; ds.

XX Montastraea cavernosa.

XX Key Location/Qualifiers

FT CDS 38..742

FT /tag= a

FT /product= "mcaVgFP protein"

PN WO2003042401-A2.

XX 22-MAY-2003.

XX 12-NOV-2002; 2002WO-US036499.

XX 13-NOV-2001; 2001US-0332980P.

XX (CLON-) CLONTECH LAB INC.

XX Labas YA, Gurskaya NG, Yanushevich Y, Fradkov AP, Lukyanov K;

PI Lukyanov S, Matz MV;

XX WPI; 2003-482353/45.

DR P-PSDB; AAO29590.

XX New chromo and/or fluorescent-protein encoding nucleic acids, the  
proteins being useful as labels in analyte detection assays, as  
selectable markers in recombinant DNA application, or as biosensors in  
prokaryotic and eukaryotic cells.

XX Claim 1; Fig 20; 89pp; English.

XX The invention relates to novel chromo and/or fluorescent proteins and



XX	SQ	Sequence	1133 BP; 344 A; 214 C; 277 G; 296 T; 0 U; 2 Other;
		Query Match	62.4%; Score 665; DB 8; Length 1133;
		Best Local Similarity	80.4%; Pred. No. 4.6e-191;
		Matches	855, Conservative 2; Mismatches 162; Indels 44; Gaps 5;
Qy	1	ATTCGCCCTCGTGATTTTGGAAAGAGAGACGATCGAGAAACAACAGAGCTGTA--AGGTTGA	58
Db	99	ATTCACCCCTCGTGATTTTGGAAAGAGAGAGATCGAGAAACAACAGAGCTGATTACGCTAA	158
Qy	59	TATCTTACTTACGCTACCATCATGACAAGTGTTCACAGAAAAGGGTGTGATTAAACC	118
Db	159	AATCTTACTTGCCTCTACCCACATGA-----GTATGATTAAACC	197
Qy	119	AGACATGAAGATGAAGCTGCGTATGGAAGGTGCTGTAAACGGGCACAAAGTTTCGTGGTTGA	178
Db	198	AGAAATGAAGATCAAGATGCGTATGGACGGTGTGTAAACGGGCACAAAGTTTCGTGATTAC	257
Qy	179	AGGAGATGGAAAAAGGGAAGCCTTTTCGACGGGAACAACAGACTATGGAACCTTACAGTCATAGA	238
Db	258	AGGGGAAGGAAGCGCGGAGCCTTTTCGAGGGAAAAACAGACTATGGAACCTGACAGTCATAGA	317
Qy	239	AGGGCACCATTTGCTTTTCGCTTACGATATCTTCGACCAACAGTATTCGATTTACGCCACAG	298
Db	318	CGCGGACCTCTGCTCTTTTCGCTTTCGACATCTTCGACCAACAGATTCGATTTACGCCACAG	377
Qy	299	GGTATTCCGCCAAATACCAGAAACATAGCAGATATTTCGAGCAGACGTTTCTCTGAGGG	358
Db	378	GGTATTCCGCCAAATACCAGAAACATCCAGAACATCCAGACTATTTCCAGCAGTCGTTTCTCTGAGG	437
Qy	359	GTACTTCTGGGAAACGAAGCATGACATACGAAGACGAGGGCATTTGCAATGCCCAACAAACGA	418
Db	438	GTTTTCTTGGGAAACGAAGCATGACTTACGAAGACGAGGGCATTTGCAATGCCCAACAAATGA	497
Qy	419	CATAAACAATGATGGAAGCGTCGACGACTGTTTTCGCTATATAAATTCGATTTGATGCTGT	478
Db	498	CATAAAATG-----GAAAGCGCACTGCTTTTCTCTATGAAATTCGATTTGATGCGGGT	548
Qy	479	GAACCTTCTCGCCAATGGTCAGATTATGCAGAGGAAGACGCTGAAATGGGAGCCATCCAC	538
Db	549	GAACCTTCTCGCCAATGATCGATTATGCAGAGGAAGACCGTGAATGGGAGCCATCCAC	608
Qy	539	TGAGATAATGTATCGCGTGATGAGGTCTCGAAGGGTGATGTTAAACATGCTCTGTTGCT	598
Db	609	TGRGGAATGTATGCGGTGATGAGGTGCTTAAAGGTGGTCTTAAACATGGCTCTGTTGCT	668
Qy	599	TGAAGGAGTGGCCATTACCGATGTGACTTCAAAACTACTTACAAAGCTTAAGAGGTTGT	658
Db	669	TGAAGGAGTGGCCATTTCGATGTGACTTGAAAACTACTTACAAAGCTTAAGAGGTTGT	728
Qy	659	CCGTTCCAGACTATCACTTTGTGGACCATCGCATTTGAGATTTGTAGCGCACGACAAAGA	718
Db	729	CCAGATGCCAGACTATCACTTTGTGAATCACCGACTTGAGATAACATGGCATGACGAGGA	788
Qy	719	TTACAACAAGGTTAAGCTGCAACGAGCATGCCGAAGCTCGTCATGGACTGTCAAGGAAGGC	778
Db	789	TTACAACAAGGTTAAGCTGTCTGAGCATGCAGAAAGCTCATTTCTGCACTGCCAAGGCAAGC	848
Qy	779	CAAGTAAAGGCTTAATGAAAGTCAAGACGCAACCGAGGAGAAACA-----AAGTA	829
Db	849	CAAAATAAAGGCTTGACGAAAAAGCCAAACGGCAAGAGTACAAGAAAGTATATATAAATG	908
Qy	830	CTTTTGTGTTAAATTTGAAGCAATTTACTCGGAATTAGTATTTTCATCTTTCGATTCGAAG	889
Db	909	TATATTTTTCAACTGAAAGGCCATTTCCACTCGGAATTAGTATTTTCATCTTTCATATTCAG	968
Qy	890	GATTTGTTCCGGATTTGTTAGAGACTAGCTCTAGAGTTGATTTTGTGAAAAAAGATAG	949
Db	969	GATTTATTTGGGATTTGCTAGCCACTAGCTTTATTTGTTAAATTAAGTTAAAGACG---G	1025
Qy	950	TTTCCAGTTTTTTCGGGATTTACAGATGGGGATAGCTTTTAACTCAGTTGTGGTCAA	1009

Db	1026	TTTAGCAATTTTTCGGTATTACAACTAGGCACAGAGCTCTTAACCCAGTAGTGGTCAG	1085
Qy	1010	ATGCAAGTAAGAAAACCTGTAGTGAGATAAACTTGTATTATCGAA	1052
Db	1086	GTACAAGTAAGAAAACCTTGGTGAGATAAGACTTGTAGTCGAA	1128
RESULT 4			
ADH00980			
ID	ADH00980	standard; DNA; 694 BP.	
XX	AC	ADH00980;	
XX			
DT	11-MAR-2004	(first entry)	
XX		Montastraea cavernosa green fluorescent protein, Green II DNA.	
DE	XX		
XX		Transcription regulation; biological study; drug screening;	
KW	XX	internal ribosome entry site; IRES;	
KW	XX	transcription-translation coupled system; green fluorescent protein;	
KW	XX	Green II; gene; ds.	
OS	XX	Montastraea cavernosa.	
XX			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..684	
FT		/*tag= a	
FT		/product= "Green II protein"	
XX			
PN	US2003157643-A1.		
XX			
PD	21-AUG-2003.		
XX			
PF	09-DEC-2002; 2002US-00314827.		
XX			
PR	24-AUG-2000; 2000US-00645706.		
XX			
PA	(ALMO//	ALMOND B D.	
PA	(WOOD//	WOOD M G.	
PA	(WOOD//	WOOD K V.	
XX			
PI	Almond BD,	Wood MG, Wood KV;	
XX			
XX	WPI; 2003-830146/77.		
DR	P-PSDS; ADH00981.		
XX			
PT	Novel synthetic nucleic acid molecule from an aquatic species useful for		
PT	imaging systems for in vivo biological studies or drug screening,		
PT	comprises nucleotides of a coding region for a fluorescent polypeptide.		
XX			
PS	Disclosure; SEQ ID NO 21; 54pp; English.		
XX			
CC	The invention relates to a synthetic nucleic acid molecule comprising		
CC	nucleotides of a coding region for a fluorescent polypeptide having a		
CC	codon composition differing at more than 25% of codons from a parent		
CC	nucleic acid sequence encoding a fluorescent polypeptide, where the		
CC	nucleic acid has at least 3-fold fewer transcription regulatory sequences		
CC	relative to average number of such sequences in the parent nucleic acid		
CC	sequence. The invention is useful for in vivo biological studies or drug		
CC	screening and is useful for detection of rare events that require extreme		
CC	sensitivity (e.g. studying RNA encoding), in use with internal ribosome		
CC	entry sites (IRES), to improve the efficiency of in vitro translation or		
CC	in vitro transcription-translation coupled systems such as TNT (RTM),		
CC	study of fluorescent proteins optimised to different host organisms (e.g.		
CC	plants, fungi). The synthetic fluorescent proteins is used as reporters		
CC	in multiwell assays and drug screening. The invention exhibits		
CC	significantly enhanced levels of mammalian expression without negatively		
CC	affecting other desirable physical or biochemical properties and has		
CC	greatly reduced number of known transcription regulatory sequences. The		
CC	present sequence is Montastraea cavernosa green fluorescent protein,		
CC	Green II DNA.		
XX			
XX	Sequence 684 BP; 211 A; 137 C; 176 G; 160 T; 0 U; 0 Other;		
XX			

Query Match	58.9%;	Score 628;	DB 10;	Length 684;
Best Local Similarity	94.9%;	Pred. No. 6.4e-180;		
Matches 649;	Conservative	0;	Mismatches 35;	Indels 0;
Gaps	0;			
QY 102	AAGGCTGTGATTTAAAC	CAGACATGAAGATGAAGCTGCGTATGGAAGGTGCTGTAAACGGG	161	
DB				
QY 1	ATCAGTGTGATAAATAC	CAGACATGAAGATCAAGCTGCGTATGGAAGGTGCTGTAAACGGG	60	
DB				
QY 162	CACAAGTTCTGCTGTTCAAG	CGAGATGAAAGGAAGGACCTTTTCGACGGGAACACAGACTATG	221	
DB				
QY 61	CACAAGTTCTGCTGTTCAAG	CGAGATGAAAGGAAGGACCTTTTCGACGGGAACACAGACTATG	120	
DB				
QY 222	GACCTTACAGTCATAGAAG	CGCACCATTTGCTTTACGATATCTTGACAAACAGTA	281	
DB				
QY 121	GACCTTACAGTCATAGAAG	CGCACCATTTGCTTTACGATATCTTGACAAACAGTA	180	
DB				
QY 282	TTTCGATTACGGCAACAGG	GATTTCCGCCAATACCCAGAAGACATAGCAGATTATTTCAAG	341	
DB				
QY 181	TTTCGATTACGGCAACAGG	GATTTCCGCCAATACCCAAAAGACATACCCAGACTATTTCAAG	240	
DB				
QY 342	CAGACGTTTCTTGAGGGG	TACTTCTGGGAAACGAAGCATGACATACGAAGACCCAGGGCATT	401	
DB				
QY 241	CAGACGTTTCTTGAGGGG	TACTTCTGGGAAACGAAGCATGACATACGAAGACCCAGGGCATT	300	
DB				
QY 402	TGCATTCGCCCAAAACGA	CATTAACAATGATGAAGCGTCGACGACTGTTTGGCTATATAA	461	
DB				
QY 301	TGCATTCGCCCAAAACGA	CATTAACAATGATGAAGCGTCGACGACTGTTTGGCTATATAA	360	
DB				
QY 462	ATTTCGATTGTATGGTGTC	AACTTTCTGCCAATGGTCAGTTATGCAGAGGAAGACGGTG	521	
DB				
QY 361	ATTTCGATTGTATGGTGTC	AACTTTCTGCCAATGGTCAGTTATGCAGAGGAAGACGGTG	420	
DB				
QY 522	AAATGGGAGCCATCCACT	CGACATATGTATCGCGGTGATGAGTGTCTGAAGGGTGATGTT	581	
DB				
QY 421	AAATGGGAGCCATCCACT	CGACATATGTATGTGTGTGTATGGGTACTGTAAGGGTGATGTT	480	
DB				
QY 582	AACATGGCTCTGTTCCT	TGAAGAGGTGGCCATTAACGATGTGACTTTCAAACTACTTAC	641	
DB				
QY 481	AACATGGCTCTGTTCCT	TGAAGAGGTGGCCATTAACGATGTGACTTTCAAACTACTTAC	540	
DB				
QY 642	AAAGCTAAGAAGGTTGTTC	CGGTTGCCAGACTATCACTTTGTGACCATCGCATTGAGATT	701	
DB				
QY 541	AAAGCTAAGAAGGTTGTTC	CGGTTGCCAGACTATCATTTTGTGACCATCGCATTGAGATT	600	
DB				
QY 702	GTGAGCCACGAACAAAG	ATTACAAAGGTTAAAGTGTGACGAGCATGCCAAGCTGTCAT	761	
DB				
QY 601	GTGAGCCACGACAAAG	ATTACAAAGGTTAAAGTGTGACGAGCATGCCAAGCTGTCAT	660	
DB				
QY 762	GGACTGTCAAGGAAGGC	CAAGTAA 785		
DB				
QY 661	GGGCTGCCGAGGAGGC	CAAGTAA 684		
DB				

RESULT 5	
ADO78073	
ID	ADO78073 standard; DNA; 684 BP.
XX	
XX	
AC	ADO78073;
XX	
XX	
DT	26-AUG-2004 (first entry)
XX	
DE	M. cavernosa green fluorescent protein mutant GreenII DNA.
XX	
KW	fluorescent protein; red fluorescent protein; green fluorescent protein;
KW	gene expression detection; regeneration; multiple labeling system;
KW	in vivo marker; microinjection assay; taxonomic marker; colour indicator;
KW	food additive; cosmetic; gene; ds; mutant; GreenII; great star coral.
OS	Montastraea cavernosa.
XX	
Key	Location/Qualifiers
FH	1. .684
FT	CDS

FT	/tag= a
FT	/product= "M. cavernosa green fluorescent protein mutant
FT	GreenII"
XX	
PN	US2004110225-A1.
XX	
PD	10-JUN-2004.
XX	
PF	09-DEC-2002; 2002US-00314936.
XX	
PR	09-DEC-2002; 2002US-00314936.
XX	
PA	(GIBB/) GIBBS P D L.
PA	(CART/) CARTER R W.
PA	(SCHM/) SCHMALE M C.
XX	
PI	Gibbs PDL, Carter RW, Schmale MC;
XX	
DR	WPI; 2004-467659/44.
DR	P-PSDB; ADO78074.
XX	
PT	New first mutant fluorescent protein having an optical property
PT	relatively different to a corresponding optical property produced by a
PT	wild-type red or green fluorescent protein, useful for detecting
PT	expression of a gene.
XX	
PS	Claim 33; SEQ ID NO 7; 30pp; English.
XX	
CC	The invention describes a first mutant fluorescent protein (I) comprising
CC	an amino acid sequence that differs from that of a corresponding wild-
CC	type red fluorescent protein by an amino acid substitution, where the
CC	first mutation fluorescent protein mutant has an optical property that
CC	differs relative to the corresponding optical property produced by the
CC	wild-type red or green fluorescent protein. (I) is useful for detecting
CC	expression of gene. The polynucleotide (II) encoding (I) is useful for
CC	detecting expression of a gene which involves introducing (II) into a
CC	cell or organism, allowing the cell to replicate, and detecting
CC	expression of the nucleic acid by emission of fluorescent light. The
CC	expression of the nucleic acid expression is detected in vivo or in
CC	vitro. (II) is useful for regenerating an animal such as a zebrafish
CC	which contains (II). (I) is useful as a marker for detecting expression
CC	of a gene, in biochemical assays, and as reagents. (I) is also useful in
CC	multiple labelling systems, as in vivo markers such as in mRNA
CC	microinjection assays, and as taxonomic markers for studies of genetics,
CC	colour indicators in diagnostic kits, coloured food additives, and
CC	cosmetics ingredients. (I) is useful in research for up or down
CC	regulation, to monitor promoter activity, to allow longer term monitoring
CC	and to localise proteins. (I) has enhanced properties such as
CC	substantially enhanced fluorescence and reduced toxicity. This sequence
CC	encodes GreenII, a mutant of the wild type green fluorescent protein
CC	isolated from great star coral.
XX	
SQ	Sequence 684 BP; 211 A; 137 C; 176 G; 160 T; 0 U; 0 Other;

	Query Match	58.9%;	Score 628;	DB 12;	Length 684;
	Best Local Similarity	94.9%;	Pred. No. 6.4e-180;		
	Matches 649;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;
Qy	102	AAGGGTGTGATTAACACGACATGAAGATGAAGCTCGGTATGCAAGGTGCTGTAAACGGG	161		
Db	1	ATGAGTGTGATAAACACGACATGAAGATCAAGCTCGGTATGGAAGTGTGTAAACGGG	60		
Qy	162	CACAGTTTCGTGTTGAAGAGATGGAAGAAGGAAGCCTTTTCGACGGAAACACAGACTATG	221		
Db	61	CACAGTTTCGTGATTGAAGGAGACGGAAGAAGGCAAGCCTTTCGAGGGAAACACAGACTATG	120		
Qy	222	GACCTTTACAGTCATAGAAGCGCGACCATTTGCTTTTCGTATTACGATATCTTTGACAAACAGTA	281		
Db	121	GACCTTTACAGTCATAGAAGCGCACCTTTGCTTTTCGCTTTCGATATCTTTGACACAGTA	180		
Qy	282	TTTCGATTACGGCAACAGGGTATTTCGCCAAATATCCCGAAGAGACATACGAGATTATTTCAAG	341		
Db	181	TTTCGATTACGGCAACAGGGTATTTCGCCAAATATCCCGAAGAGACATACGAGACTATTTCAAG	240		



QY 342 CAGACGTTTCTGAGGGTACTTCTGGAAACGAAGCATGACATACAGACAGGCGATT 401  
Db 241 CAGACGTTTCTGAGGGGACTCTCTGGAAACGAAGCATGACATACAGACAGGCGATT 300  
QY 402 TGCATCGCCACAAACGACATAACAATGATGGAAGGCGTCGACGACTGTTTTCCTATATAA 461  
Db 301 TGCATCGCCACAAACGACATAACAATGATGGAAGGCGTCGACGACTGTTTTCCTATATAA 360  
QY 462 ATTTCGATTTGATGGTGTGAATCTTCTGCAATGTCCTGATGTCAGAGGAGAGCGTG 521  
Db 361 ATTTCGATTTGATGGTGTGAATCTTCTGCAATGTCCTGATGTCAGAGGAGAGCGTGA 420  
QY 522 AAATGGAGCGCATCCACTGAGATAATGATGCGCGTGAAGGCTGCTGAAGGCTGATGTT 581  
Db 421 AAATGGAGCGCATCCACTGAGATAATGATGCGCGTGAAGGCTGCTGAAGGCTGATGTT 480  
QY 582 AACATGGCTCTGTTGCTTGAAGGAGGTGGCCATTACCGATGTGACATTTCAAAACTACTTAC 641  
Db 481 AACATGGCTCTGTTGCTTGAAGGAGGTGGCCATTACCGATGTGACATTTCAAAACTACTTAC 540  
QY 642 AAAGCTAAGAGGTTGTCGGTTGCGAGACTATCACTTTTGTGACCATCGCATTTGAGATT 701  
Db 541 AAAGCTAAGAGGTTGTCGGTTGCGAGACTATCACTTTTGTGACCATCGCATTTGAGATT 600  
QY 702 GTGAGCCGACGACAAATACACAGGTTAAGCTGACGAGCATGCCGAGCTGCTCAT 761  
Db 601 GTGAGCCGACGACAAATACACAGGTTAAGCTGATGATGATGCCGAGCATGCCGAGCTCAT 660  
QY 762 GGACTGTCAAGGAAGGCCAAGTAA 785  
Db 661 GGCTGCCGAGGCGGCCAAGTAA 684

RESULT 6  
AD284228  
ID AD284228 standard; DNA; 684 BP.  
AC AD284228;  
XX  
XX 14-JUL-2005 (first entry)  
XX  
XX Montastraea cavernosa GFP mutant Green II-encoding DNA, SEQ ID NO:7.  
DE  
XX Green fluorescent protein; Green II; mutant; gene; ds.  
XX  
XX Montastraea cavernosa.  
OS  
OS Synthetic.  
PH  
XX Key Location/Qualifiers  
CDS 1..684  
FT /\*tag= a  
FT /product= "Mutant GFP Green II"  
FT replace(527,C)  
FT /\*tag= b  
FT mutation  
FT /note= "This nucleotide is C in the Green I mutant,  
FT resulting in Ser rather than Phe at position 176 of the  
FT protein"  
XX  
XX US2005100954-A1.  
PN  
XX  
XX 12-MAY-2005.  
PD  
XX  
XX 23-DEC-2004; 2004US-00021014.  
PF  
XX  
XX 09-DEC-2002; 2002US-00314936.  
PR  
XX  
XX (GIBB/) GIBBS P D L.  
PA (CART/) CARTER R W.  
PA (SCHW/) SCHWALE M C.  
XX  
XX Gibbs PDL, Carter RW, Schmale MC;  
PI  
XX

DR WPI; 2005-345399/35.  
DR P-PSDB; ADZ84229.  
XX  
PT Novel mutant fluorescent protein comprising amino acid sequence that  
PT differs from wild-type red or green fluorescent protein by amino acid  
PT substitutions, useful as markers for detecting desired gene expression.  
XX  
PS Example 4; SEQ ID NO 7; 29pp; English.  
XX  
CC The invention relates to mutant red and green fluorescent proteins which  
CC have higher fluorescent intensities compared to the wild-type proteins  
CC from which they are derived. The invention also discloses nucleic acids  
CC encoding the mutant fluorescent proteins, and transgenic animals  
CC comprising such a nucleic acid. The mutant red fluorescent proteins (RFP)  
CC of the invention, Red I (ADZ84223) and Red II (ADZ84225), are derived  
CC from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a  
CC mushroom coral believed to be either an Actinodiscus or Discosoma  
CC species, while the mutant green fluorescent proteins (GFP), Green I  
CC (ADZ84227) and Green (ADZ84229), are derived from a novel wild-type GFP  
CC (known as McGFP) isolated from the great star coral Montastrea  
CC cavernosa. The Ac/DsRFP and McGFP mutants were obtained by subjecting the  
CC wild-type Ac/DsRFP and McGFP cDNAs to low-stringency PCR to introduce  
CC random mutations, cloning the PCR products into bacterial expression  
CC vectors, and selecting the transformed bacteria having higher relative  
CC fluorescence compared to those expressing the wild-type proteins. The  
CC mutant fluorescent proteins of the invention may be used in a wide range  
CC of applications. They are useful as markers for detecting the expression  
CC of a gene of interest, or can be used as in vivo markers in mRNA  
CC microinjection assays in transgenic animals. They can be used in  
CC expression systems (e.g., as components of fusion proteins), in multiple  
CC labeling systems, or can be used in biochemical assays and as reagents.  
CC They may additionally be used as taxonomic markers for studies of  
CC cnidarian genetics, as color indicators in diagnostic kits, as colored  
CC food additives and as cosmetic ingredients. The present sequence  
CC represents DNA encoding the mutant Montastrea cavernosa green  
CC fluorescent protein Green II. Green II, in contrast to the mutant GFP  
CC Green I (ADZ84227), has a high resistance to photobleaching.  
XX  
SQ Sequence 684 BP; 211 A; 137 C; 176 G; 160 T; 0 U; 0 Other;

Query Match 58.9%; Score 628; DB 14; Length 684;  
Best Local Similarity 94.9%; Pred. No. 6.4e-180;  
Matches 649; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 102 AAGGTGTGATTAACCAAGACATCAAGATGAAGTGGCTGCTATGAGGTGCTGTAACCGG 161  
Db 1 ATGAGTGTGATAAAACCAAGACATCAAGATCAAGTGGCTGATGGAAGGTGCTGTAACCGG 60  
QY 162 CACAAGTTCGTGGTTGAAGGAGATGGAAGGAGGAGCCCTTCGACGGAACACAGACTATG 221  
Db 61 CACAAGTTCGTGGTTGAAGGAGAGCGGAAGGAGCCCTTCGAGGGAACACAGACTATG 120  
QY 222 GACCTTACAGTCATAGAAGGCGCACCATTTGCTTTCGCTTACGATATCTTGACACAGTA 281  
Db 121 GACCTTACAGTCATAGAAGGCGCACCTTTGCTTTCGCTTACGATATCTTGACACAGTA 180  
QY 282 TTCGATTCGCGACACAGGCTATTCGCCAAATACCCAGAGACATAGCAGATTTATTTCAAG 341  
Db 181 TTCGATTCGCGCAACAGGCTATTCGCCAAATACCCAAATACCCAAAGACATACAGACTTTTCAAG 240  
QY 342 CAGACGTTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATACAGACAGGCGCAT 401  
Db 241 CAGACGTTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATACAGACAGGCGCAT 300  
QY 402 TGCATCGCCACAAACGACATAACAATGATGGAAGGCGTCGACGACTGTTTTCCTATATAA 461  
Db 301 TGCATCGCCACAAACGACATAACAATGATGGAAGGCGTCGACGACTGTTTTCCTATATAA 360  
QY 462 ATTTCGATTTGATGGTGTGAATCTTCTGCAATGTCCTGATGTCAGAGGAGAGCGTG 521  
Db 361 ATTTCGATTTGATGGTGTGAATCTTCTGCAATGTCCTGATGTCAGAGGAGAGCGTGA 420  
QY 522 AAATGGAGCGCATCCACTGAGATAATGATGCGCGTGAAGGCTGCTGAAGGCTGATGTT 581







CC of the invention.  
 XX Sequence 684 BP; 205 A; 130 C; 181 G; 168 T; 0 U; 0 Other;  
 SQ Query Match 50.7%; Score 540; DB 14; Length 684;  
 Best Local Similarity 86.8%; Pred. No. 3.6e-153;  
 Matches 594; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 102 AAGGTTGTTGATTAAACAGACATGAAGATGAAGTGGTATGAAGTGGTCTGTAACCGG 161  
 DB 1 ATGAGTGTGATTACATCAGAAATGAAGATGAGCTGCTTATGAAGCGCTGTAACCGG 60

QY 162 CACAAGTTCCTGTTGAAGCAGATGGAAGGAGGAGCTTTCAGCGAACACAGACTATG 221  
 DB 61 CACAAGTTCCTGTTGAAGCAGATGGAAGGAGGAGCTTTCAGCGAACACAGACTATG 120

QY 222 GACCTTACAGTCTATAGAAAGCGCACCATTCCTTTCGCTTACGATATCTTGAACACAGTA 281  
 DB 121 GACCTTACAGTCTATAGAAAGCGCACCATTCCTTTCGCTTACGATATCTTGAACACAGTA 180

QY 282 TTCGATTACGGCACAGAGGTATTCGCCAATACCCAGACATACGAGATTTTCAAG 341  
 DB 181 TTCGATTACGGCACAGAGGTATTCGCCAATACCCAGACATACGAGATTTTCAAG 240

QY 342 CAGACGTTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATACGAAGCCAGGGGCAATT 401  
 DB 241 CAGTCTGTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATACGAAGCCAGGGGCAATT 300

QY 402 TGCATCGCCACAAACAGACATGAAGATGAAGTGGTATGAAGTGGTCTGTAACCGG 461  
 DB 301 TGCATCGCCACAAACAGACATGAAGATGAAGTGGTATGAAGTGGTCTGTAACCGG 360

QY 462 ATTTCGATTTGATGTTGTAAGTTCCTGCGCAATGGTCCAGTTATGAGAGGAGACCGCTG 521  
 DB 361 ATTTCGATTTGATGTTGTAAGTTCCTGCGCAATGGTCCAGTTATGAGAGGAGACCGCTG 420

QY 522 AAATGGGAGCCATCCACTGAGATAATGTATGCGCGTGAATGAGAGTGGTCTGTAACCGG 581  
 DB 421 AAATGGGAGCCATCCACTGAGATAATGTATGCGCGTGAATGAGAGTGGTCTGTAACCGG 480

QY 582 AACATGGCTCTGTTGCTTGAAGGAGTGGCCATTCACGATGTGACATTCACAAATCTTAC 641  
 DB 481 AACATGGCTCTGTTGCTTGAAGGAGTGGCCATTCACGATGTGACATTCACAAATCTTAC 540

QY 642 AAAGCTAAGAGGTTGTCGGTTCGACAGTATCATTCTTGGACCATTCGATTCAGATT 701  
 DB 541 AAAGCTAAGAGGTTGTCGGTTCGACAGTATCATTCTTGGACCATTCGATTCAGATT 600

QY 702 GTGAGCCACACAGAGATTACAAAGTTAAAGCTGCACGAGCATGCCGAAGCTCGTCAAT 761  
 DB 601 ACAAGCCATGACAGGATTACAAAGTTAAAGCTGCACGAGCATGCCGAAGCTCGTCAAT 660

QY 762 GGACTGTCAAGGAAGGCCAAGTAA 785  
 DB 661 GGCTGCTCAAGGCTGGCCNAGTAA 684'

RESULT 11  
 ID ADW43035  
 XX ADW43035 standard; DNA; 684 BP.  
 AC ADW43035;  
 XX 24-MAR-2005 (first entry)  
 DT Favia favus fluorescence-related H8pv DNA - SEQ ID 15.  
 DE Favia favus fluorescence-related H8pv DNA - SEQ ID 15.  
 XX fluorescence; protein localization; cellular transport; ds; gene.  
 KW Favia favus.  
 XX Favia favus.  
 OS Favia favus.  
 PH Key: Location/Qualifiers  
 FT CDS 1..684

FT /\*tag= a  
 FT /product= "Favia favus fluorescence-related H8pv protein  
 XX - SEQ ID 14"  
 PN WO2004111236-A1.  
 XX 23-DEC-2004.  
 XX 16-JUN-2004; 2004WO-JP008790.  
 XX 16-JUN-2003; 2003JP-00170330.  
 XX (RIKE ) RIKEN KK.  
 PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.  
 XX Miyawaki A, Tsutsumi H, Karasawa S;  
 PI WPI; 2005-066240/07.  
 DR P-PSDB; ADW43034.  
 XX Novel fluorescent protein derived from Favia favus, exhibiting specific  
 PT excitation maximum wavelength and molar extinction coefficient, useful in  
 PT molecular-biological analysis.  
 XX Claim 11; SEQ ID NO 15; 66pp; Japanese.  
 CC The invention relates to a novel fluorescent protein derived from Favia  
 CC favus, exhibiting an excitation maximum wavelength of 507 nm and  
 CC fluorescence maximum wavelength of 517 nm, with molar extinction  
 CC coefficient of 80000 at 482 nm, quantum yield of 0.68 and pH sensitivity  
 CC of fluorescence maximum being stable at pH 5-11. The protein has a fully  
 CC defined sequence of SEQ ID NO:1 as given in the specification. The  
 CC fluorescent fusion protein of the invention may be useful for analyzing  
 CC the location of transport of a protein inside the cell. The fluorescent  
 CC protein does not exhibit toxicity in a mammalian cell. The current  
 CC sequence is that of the Favia favus fluorescence-related H8pv DNA (SEQ ID  
 CC 15) of the invention.  
 XX Sequence 684 BP; 204 A; 131 C; 183 G; 166 T; 0 U; 0 Other;  
 SQ Query Match 50.4%; Score 536.8; DB 14; Length 684;  
 Best Local Similarity 86.5%; Pred. No. 3.4e-152;  
 Matches 592; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 102 AAGGTTGTTGATTAAACAGACATGAAGATGAAGTGGTATGAAGTGGTCTGTAACCGG 161  
 DB 1 ATGAGTGTGATTACATCAGAAATGAAGATGAGCTGCTGTAACCGG 60

QY 162 CACAAGTTCCTGTTGAAGGAGATGGAAGGAGGAGCTTTCGACCGAACACAGACTATG 221  
 DB 61 CACAAGTTCCTGTTGAAGGAGATGGAAGGAGGAGCTTTCGACCGAACACAGACTATG 120

QY 222 GACCTTACAGTCTATAGAAAGCGCACCATTCCTTTCGCTTACGATATCTTGAACACAGTA 281  
 DB 121 GACCTTACAGTCTATAGAAAGCGCACCATTCCTTTCGCTTACGATATCTTGAACACAGTA 180

QY 282 TTCGATTACGGCACAGAGGTATTCGCCAATACCCAGACATACGAGATTTTCAAG 341  
 DB 181 TTCGATTACGGCACAGAGGTATTCGCCAATACCCAGACATACGAGATTTTCAAG 240

QY 342 CAGACGTTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATACGAAGCCAGGGGCAATT 401  
 DB 241 CAGTCTGTTCTGAGGGGTACTTCTGGGAACGAAGCATGACATACGAAGCCAGGGGCAATT 300

QY 402 TGCATCGCCACAAACAGACATGAAGATGAAGTGGTATGAAGTGGTCTGTAACCGG 461  
 DB 301 TGCATCGCCACAAACAGACATGAAGATGAAGTGGTATGAAGTGGTCTGTAACCGG 360

QY 462 ATTTCGATTTGATGTTGTAAGTTCCTGCGCAATGGTCCAGTTATGAGAGGAGACCGCTG 521  
 DB 361 ATTTCGATTTGATGTTGTAAGTTCCTGCGCAATGGTCCAGTTATGAGAGGAGACCGCTG 420

QY 522 AAATGGGAGCCATCCACTGAGATAATGTATGCGCGTGAATGAGAGTGGTCTGTAACCGG 581

Db 421 AAATGGAGCCATCCACTGAGAAATGTATGTCGTGATGAGTGTGCTGAAGGCTGATGTT 480  
 QY 582 AACATGGCTCTGTGCTTCAAGAGGTGGCCATTACCGATGTGACTTCAAACTACTTAC 641  
 Db 481 AACATGGCTCTGTGCTTCAAGAGGTGGCCATTACCGATGTGACTTCAAACTACTTAC 540  
 QY 642 AAAGCTAAGAGGTTGTCCGTTGCCAGACTATCACTTTGTGGACCATCGCATTGAGATT 701  
 Db 541 AAAGCAAGAGGTTGTCCGTTGCCAGACTATCACTTTGTGGATCATCGATGAGATA 600  
 QY 702 GTGAGCCAGCAAGATTACCAAGGTTAAGCTGCGAGCATGCCGAAGCTCGTCAAT 761  
 Db 601 ACAAGCCATGACAAGGATTACCAAGGTTAAGCTGATGAGCATGCTAAAGCTCATTC 660  
 QY 762 GGACTGTCAAGGAAGGCCAAGTAA 785  
 Db 661 GGGCTGCCAAGGCTGGCCCAAGTAA 684

RESULT 12  
 ADW43041  
 ID ADW43041 standard; DNA; 684 BP.  
 AC ADW43041;  
 XX  
 DT 24-MAR-2005 (first entry)  
 XX  
 DE Favia favus fluorescence-related KBL2 DNA - SEQ ID 21.  
 XX  
 KW fluorescence; protein localization; cellular transport; ds; gene.  
 XX  
 OS Favia favus.  
 XX  
 Key Location/Qualifiers  
 CDS 1..684  
 /tag= a  
 /product= "Favia favus fluorescence-related KBL2 protein  
 - SEQ ID 21"  
 /transl\_except= (pos:280..282, aa:Xaa)

WO2004111236-A1.  
 XX  
 PN 23-DEC-2004.  
 XX  
 PD 16-JUN-2004; 2004WO-JP008790.  
 XX  
 PF 16-JUN-2003; 2003JP-00170330.  
 XX  
 PR (RIKE ) RIKEN KK.  
 PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.  
 XX  
 PI Miyawaki A, Tsutsui H, Karasawa S;  
 XX  
 DR WPI; 2005-066240/07.  
 DR P-PSDB; ADW43040.  
 XX

Novel fluorescent protein derived from Favia favus, exhibiting specific excitation maximum wavelength and molar extinction coefficient, useful in molecular-biological analysis.  
 Claim 11; SEQ ID NO 21; 66pp; Japanese.  
 The invention relates to a novel fluorescent protein derived from Favia favus, exhibiting an excitation maximum wavelength of 507 nm and fluorescence maximum wavelength of 517 nm, with molar extinction coefficient of 80000 at 482 nm, quantum yield of 0.68 and pH sensitivity of fluorescence maximum being stable at pH 5-11. The protein has a fully defined sequence of SEQ ID NO.1 as given in the specification. The fluorescent fusion protein of the invention may be useful for analyzing the location or transport of a protein inside the cell. The fluorescent protein does not exhibit toxicity in a mammalian cell. The current sequence is that of the Favia favus fluorescence-related KBL2 DNA (SEQ ID

CC 21) of the invention.  
 XX  
 SQ Sequence 684 BP; 203 A; 133 C; 183 G; 165 T; 0 U; 0 Other;  
 Query Match 50.1%; Score 533.6; DB 14; Length 684;  
 Best Local Similarity 86.3%; Pred. No. 3.2e-151;  
 Matches 590; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
 QY 102 AAGGTGTGATTAACACAGACATCAAGATGAAGCTGCGTATGGAAGTCTGTAAACGGG 161  
 Db 1 ATGAGTGTGATTAATCATCAGAAATGAGATGGAGCTGGTATGGAAGGCGCTGTAAACGGG 60  
 QY 162 CACAAGTTCGTGTTGAAGGAGATGGAAGGGAAGGCTTTCGACGGAACACAGACTATG 221  
 Db 61 CACAAGTTCGTGTTGAAGGGAAGGCTTTCGACGGAACACAGACTATG 120  
 QY 222 GACCTTACAGTCATAGAGGGGACCATTCGCTTTCGCTTACGATATCTTGACACAGTA 281  
 Db 121 GACCTTACAGTCATAGAGGGGACCATTCGCTTTCGCTTTCGATATCTTGACACAGTA 180  
 QY 282 TTCGATTACGGCAACAGGGGTATTTCGCCAAATACCCAGAGACATAGCAGATTATTTCAAG 341  
 Db 181 TTCGGTCACGGCAACCGGGTATTTCGCAATATCCAGAGAAATAGTAGACTACTTCAAG 240  
 QY 342 CAGACGTTTCTGAGGGGTAATTCTTGGGAACGAGCATGACATACGAAAGACCGGGCAT 401  
 Db 241 CAGTCGTTTCTGAGGGTATTCTTGGGAACGAGCATGAGTTACGAGACGGGGGAATT 300  
 QY 402 TGCATCGCCACAAACGACATACATGATGGAAGGCGTCGACGACTGTTTTCCTATAAA 461  
 Db 301 TGCCTCGCCACAAACATATATACGATGAAGAAAGACGGCAGCAACTGTTTGTCTATGAA 360  
 QY 462 ATTCCGATTGATGGTGTGAACCTTTCCTGCCAATGTCAGTTATGCGAGTGTGAGGAGAGCGCTG 521  
 Db 361 ATTCCGATTGATGGTGTGAACCTTTCCTGCCAATGTCAGTTATGCGAGTGTGAGGAGAGCGCTG 420  
 QY 522 AAATGGAGCCATCCACTGAGATAATGATGCGCGTGTGAGTGTGAGGAGTGTGAGGAGTGTG 581  
 Db 421 AAATGGAGCCATCCACTGAGATAATGATGCGCGTGTGAGTGTGAGGAGTGTGAGGAGTGTG 480  
 QY 582 AACATGGCTCTGTGCTTGAAGGAGTGGCCATTACCGATGTGACTTCAAACTACTTAC 641  
 Db 481 AACATGGCTCTGTGCTTCAAGAGGTGGCCATTACCGATGTGACTTCAAACTACTTAC 540  
 QY 642 AAAGCTAAGAGGTTGTCCGTTGCCAGACTATCACTTTGTGGACCATCGCATTGAGATT 701  
 Db 541 AAAGCAAGAGGTTGTCCGTTGCCAGACTATCACTTTGTGGATCATCGATGAGATA 600  
 QY 702 GTGAGCCAGCAAGATTACCAAGGTTAAGCTGCGAGCATGCCGAAGCTCGTCAAT 761  
 Db 601 ACAAGCCATGACAAGGATTACCAAGGTTAAGCTGATGAGCATGCTAAAGCTCATTC 660  
 QY 762 GGACTGTCAAGGAAGGCCAAGTAA 785  
 Db 661 GGGCTGCCAAGGCTGGCCCAAGTAA 684

RESULT 13  
 AAL60080  
 ID AAL60080 standard; DNA; 795 BP.  
 XX  
 AC AAL60080;  
 XX  
 DT 27-AUG-2003 (first entry)  
 XX  
 DE Montastrea cavernosa red fluorescent protein (mcavRFP) encoding DNA.  
 XX  
 KW Chromoprotein; fluorescent protein; colouring agent; food composition;  
 KW cosmetic; fluorescence resonance energy transfer; calcium ion indicator;  
 KW biosensor; pH indicator; phosphorylation indicator; messenger detector;  
 KW transgenic animal; fluorescent timer; red fluorescent protein; RFP;  
 KW transgenic; gene; ds.  
 XX



Novel fluorescent protein derived from *Favia fava*, exhibiting specific excitation maximum wavelength and molar extinction coefficient, useful in molecular-biological analysis.

Claim 11; SEQ ID NO 17; 66pp; Japanese.

The invention relates to a novel fluorescent protein derived from *Favia fava*, exhibiting an excitation maximum wavelength of 507 nm and fluorescence maximum wavelength of 517 nm, with molar extinction coefficient of 8000 at 482 nm, quantum yield of 0.68 and pH sensitivity of fluorescence maximum being stable at pH 5-11. The protein has a fully defined sequence of SEQ ID NO:1 as given in the specification. The fluorescent fusion protein of the invention may be useful for analyzing the location or transport of a protein inside the cell. The fluorescent protein does not exhibit toxicity in a mammalian cell. The current sequence is that of the *Favia fava* fluorescence-related H8PVLN DNA (SEQ ID 17) of the invention.

Sequence 684 BP; 205 A; 131 C; 182 G; 166 T; 0 U; 0 Other;

Query Match 49.5%; Score 527.2; DB 14; Length 684;  
Best Local Similarity 85.7%; Pred. No. 2.8e-149;  
Matches 586; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

102 AAGGGTGTGATTAAACACACATGAAGTGAAGTGGTATGAGAGTGGTCTGTAACCGG 161  
1 ATGAGTGTGATTACATCAGAAATGAAGTCGAGTGGTATGGAAGGCGCTGTAAACCGG 60  
162 CACAAATTTGTTGTTGAAGGAGATGAAGAAAGGAAAGCCCTTCGACGGAACACAGACTATG 221  
61 CACAAATTTGTTGATTACAGGGAAGAAAGTGGCCAGCCCTTCGAGGGAATACAGATGTG 120  
222 GACCTTACAGTATGAAGGCGCACATTCGCTTTCGCTTACGATATCTTGACAAAGTA 281  
121 GACCTGACAGTATGAAGGCGGAGCTCTCTCTTTTGTCTTGCATATCTGACACAGCA 180  
282 TTCGATTACGGCAACAGAGTATTCGCCAAATACCCAGAGACATAGCAGATTTTCAAG 341  
181 TTCGATTACGGCAACAGGATTTTGTCAAAATACCCAGAGAAATAGTAGACTACTTCAAG 240  
342 CAGAGCTTTCTGAGGGGTACTTCTGGGAAAGCAAGCATGACATACGAAGCAGGGCAAT 401  
241 CAGTGTCTCTGAGGGTATTTCTGGGAAAGCAAGCATGAGTACGAGAGCGGGGATTT 300  
402 TGCATCGCCACAAACGACATACAAATGATGGAAGCGGTCGACGACTGTTTTCCTTATAA 461  
301 TGCCTCGCCACAAACATATACGATGAAGAAAGACGCGCAGCAACTGTTTGTCAATGA 360  
462 ATTCGATTGATGTTGTAACCTTCTGCCAATGTCGAGTATGTCAGAGGAGACGCTG 521  
361 ATTCGATTGATGTTGTAACCTTCTGCCAATGTCGAGTATGTCAGAGGAGACGCTG 420  
522 AAATGGGAGCCATCCACTGAGATATGATGTCGCGTATGATGAGTGTCTGAAGGGTATGT 581  
421 AAATGGGAGTATCCACTGAGAAATGATGTCGCTGATGAGTGTCTGAAGGGTATGT 480  
582 AACATGGCTCTGTTGTTGAAGGAGGTGGCCATACCATGATGATCTTCAAAACTACTTAC 641  
481 AACATGGCTCTGTTGTTGAAGGAGGTGGCCATACCATGATGATCTTCAAAACTACTTAC 540  
642 AAAGCTAGAGAGGTGTCGGTTCGACATATCATCTTGTGGACCATCGCATTTGAGATT 701  
541 AAAGCAAGAGAGGTGTCGGTTCGACATATCATCTTGTGGATCATCTTAAATGAGATA 600  
702 GTGAGCCACGACAAAGATTACAAAGGTTAAGCTGACGAGCATGCGGAGGCTCGTCAT 761  
601 ACAAGCCATGACAGGATTAACAAGGTTAAGCTGTATGAGCATGCTTAAAGCTCATTC 660  
762 GGACTGTCAAGGAAGGCGCAAGTAA 785  
661 GGGCTGCCAAGGCTGGCCAAAGTAA 684

RESULT 15

ADW43039  
ID ADW43039 standard; DNA; 684 BP.

XX AC ADW43039;

XX DT 24-MAR-2005 (first entry)

XX DE *Favia fava* fluorescence-related KikGR DNA - SEQ ID 19.

XX KW fluorescence; protein localization; cellular transport; ds; gene.

XX OS *Favia fava*.

XX FH Key Location/Qualifiers

XX FT 1..684

XX FT /tag= a

XX FT /product= "Favia fava fluorescence-related KikGR protein

XX FT - SEQ ID 18"

XX PN WO2004111236-A1.

XX PD 23-DEC-2004.

XX PF 16-JUN-2004; 2004WO-JP008790.

XX PR 16-JUN-2003; 2003JP-00170330.

XX PA (RIKE) RIKEN KK.

XX PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.

XX PI Miyawaki A, Tsutsui H, Karasawa S;

XX DR WPI; 2005-066240/07.

XX DR P-PSDB; ADW43038.

XX PT Novel fluorescent protein derived from *Favia fava*, exhibiting specific

XX PT excitation maximum wavelength and molar extinction coefficient, useful in

XX PT molecular-biological analysis.

XX PS Claim 11; SEQ ID NO 19; 66pp; Japanese.

XX CC The invention relates to a novel fluorescent protein derived from *Favia*

XX CC *fava*, exhibiting an excitation maximum wavelength of 507 nm and

XX CC fluorescence maximum wavelength of 517 nm, with molar extinction

XX CC coefficient of 8000 at 482 nm, quantum yield of 0.68 and pH sensitivity

XX CC of fluorescence maximum being stable at pH 5-11. The protein has a fully

XX CC defined sequence of SEQ ID NO:1 as given in the specification. The

XX CC fluorescent fusion protein of the invention may be useful for analyzing

XX CC the location or transport of a protein inside the cell. The fluorescent

XX CC protein does not exhibit toxicity in a mammalian cell. The current

XX CC sequence is that of the *Favia fava* fluorescence-related KikGR DNA (SEQ

XX CC ID 19) of the invention.

XX SQ Sequence 684 BP; 206 A; 133 C; 182 G; 163 T; 0 U; 0 Other;

Query Match 49.5%; Score 527.2; DB 14; Length 684;

Best Local Similarity 85.7%; Pred. No. 2.8e-149;

Matches 586; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 102 AAGGGTGTGATTAAACACGACATGAAGTGAAGTGGTATGAGAGTGGTCTGTAACCGG 161

Db 1 ATGAGTGTGATTACATCAGAAATGAAGTCGAGTGGTATGGAAGGCGCTGTAAACCGG 60

Qy 162 CACAAATTTGTTGTTGAAGGAGATGAAGAAAGGAAAGCCCTTCGACGGAACACAGACTATG 221

Db 61 CACAAATTTGTTGATTACAGGGAAGAAAGTGGCCAGCCCTTCGAGGGAATACAGAAATGTG 120

Qy 222 GACCTTACAGTATGAAGGCGCACCATTCGCTTTTCGCTTACGATATCTTGACAAAGTA 281

Db 121 GACCTGACAGTATAGAGGCGGAGCCCTCTCTCTTTGCTTTGCTATCTTGACAAAGCA 180



